ARI Research Note 96-25

Occupational Analysis and Job Structures

Mary Ann Statman, Monica Gribben,
Dick A. Harris, and Gene R. Hoffman
Human Resources Research Organization

Selection and Assignment Research Unit
Michael G. Rumsey, Chief

February 1996

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Human Resources Research Organization

Technical review by

Jay M. Silva
Clinton B. Walker

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1996, February	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) September 30 1994

4. TITLE AND SUBTITLE Occupational Analysis and Job Structures	5a. CONTRACT OR GRANT NUMBER MDA903-93-C-0089 5b. PROGRAM ELEMENT NUMBER 0602785A
6. AUTHOR(S) Mary Ann Statman, Monica Gribben, Dick A. Harris, and Gene R. Hoffman	5c. PROJECT NUMBER A790 5d. TASK NUMBER 5901 5e. WORK UNIT NUMBER C06
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization (HumRRO0 66 Canal Center Plaza, Suite 400 Alexandria, VA 22314	8. PERFORMING ORGANIZATION REPORT NUMBER Final Report FR-PRD-94-28
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-RS 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 96-25

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: Gabriel P. Intano

14. ABSTRACT *(Maximum 200 words)*:

The objective of this study was to review the major issues and trends in job family research and to develop a method for evaluating the quality of job cluster structures. The proposed cluster evaluation method consists of the following four components: (1) internal validation, (2) consistency analysis, (3) external comparisons and (4) validation against an external criterion. Four Army databases containing job analysis information were used to test the evaluation method. Clusters were formed by three empirical procedures: (1) Ward hierarchical cluster analysis (HCA), (2) average linkage HCA, and (3) K-means partitional clustering. The method was useful in constructing job clusters, evaluating their consistency across clustering procedures and samples, and in making external comparisons with other job family structures. It can be applied to any cluster structure evaluation problem, specifically, in the present context, to constructing new job families, developing task clusters for structuring and restructuring jobs, and evaluating the quality of existing cluster structures.

15. SUBJECT TERMS

Job family Cluster analysis Task analysis Job analysis
Numberical taxonomy Job clustering

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	173	

FOREWORD

As part of its Innovative Ideas From Industry Program, the Selection and Assignment Research Unit (SARU) of the Manpower and Personnel Research Division (MPRD) at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) contracted with the Human Resources Research Organization to develop a multipurpose method for assessing and classifying Military Occupational Specialties (MOS) and their associated duties along a variety of dimensions. The clustering of MOS, and their duties and tasks, is an important tool for simplifying the design of personnel management processes, including selection and classification procedures. However, there is an absence of research on techniques for evaluating the quality of job structures. This report describes a method, developed and tested in this contract, for evaluating classifications created for a variety of purposes. The approach can be employed in basic and applied studies to assess the quality of job structures.

ZITA M. SIMUTIS
Deputy Director
(Science and Technology)

EDGAR M. JOHNSON
Director

ACKNOWLEDGMENTS

We would like to thank the Contracting Officer's Representative, Gabriel P. Intano, for his ideas and direction in this effort. We also would like to thank Ric Blacksten for his help in conceptualizing the problem and in searching for methods of representing job clusters, and Paul J. Sticha for his thoughts on interpreting the results.

OCCUPATIONAL ANALYSIS AND JOB STRUCTURES

EXECUTIVE SUMMARY

Research Requirement:

This project had two objectives. The first was to study the major issues and trends in job clustering research. The second was to develop a method for evaluating the quality of job family structures and for comparing alternative clusters developed for different purposes (e.g., personnel classification or career development). Phase 1 of the study provided the background for developing the cluster evaluating method. The job family and cluster analysis literatures were reviewed, and the Army job family structures, developed for both operational and research purposes, were described. The literature review found that no formal cluster validation methods exist for studying applied research questions where the true cluster structure of the data is unknown.

Phase 2 involved developing and testing a method to validate new job families and to assess the quality of existing cluster structures. The core research question underlying the method was how to reconcile the differences in the configurations of job families found in the literature and within the Army.

Procedure:

The proposed cluster evaluation method consists of a set of tools (analytic procedures and indexed) that can be used in both basic research and applied studies, and includes the following components: (1) internal validation, (2) consistency analysis, (3) external comparisons, and (4) validation against an external criterion. Internal validation of a cluster structure involves identifying the number and composition of clusters that provides the best representation of the underlying relationships among the objects. Consistency analysis is a replication method of evaluating whether the observed cluster structure is the true structure and involves comparison of groups across clustering procedures or samples. External comparisons involve evaluating the congruence of cluster solutions obtained from different sources of data. External validation, the investigation of which was beyond the scope of this study, is the assessment of the usefulness of job clusters for a specific personnel process, e.g., performance appraisal.

Four existing Army databases containing job analysis information were used to test the evaluation method. Clusters were formed by three empirical procedures: (1) Ward hierarchical cluster analysis (HCA), (2) average linkage HCA, and (3) K-means

partitional clustering. A large number of solutions were examined for each procedure in each database.

Two internal validity indexes were used to select cluster solutions that provided the best recovery of the relationships among the jobs. Two types of consistency analysis were conducted. The first compared pairs of cluster solutions produced by alternative procedures in two of the four databases. The second was a double cross-validation procedure that compared the cluster structures of cross-samples in one of the databases. The external analyses measured agreement between job families based on different types of job descriptors and between the empirical clusters and the Army Aptitude Area (AA) job families.

Findings:

The internal validation procedure selected cluster solutions that were moderately to highly consistent across cluster procedures within a particular database. Further, the results were consistent with previous research on the same databases. In contrast, the cross-validation procedure found that the clusters were quite different across samples. The discrepancies may be due to the presence of overlapping families containing jobs that fit into more than one cluster. The external comparisons showed little congruence in the clusters of different databases and in the empirical and Army AA job families.

Utilization of Findings:

The proposed method was useful in constructing job clusters, evaluating their consistency across clustering procedures and samples, and making external comparisons with other job family structures. External validation of cluster structures against a relevant criterion is also an important component of the evaluation method, but was beyond the scope of the present study. The main limitation of the method is the absence of statistical tests for all but one of the three indexes we examined. This was partially overcome by the consistency analyses. Further, two approaches were suggested for developing sampling distributions of the internal and external validity indexes.

The proposed cluster evaluation method can be applied to any cluster evaluation problem, specifically, constructing new job families, developing task clusters for structuring and restructuring jobs, and evaluating the quality of existing cluster structures. The results suggest that further research is needed to evaluate the Army AA and Career Management Fields.

OCCUPATIONAL ANALYSIS AND JOB STRUCTURES

CONTENTS

Page

CHAPTER I - INTRODUCTION 1

Review of the Job Family Research 1
Army Job Families 11
Research Questions 15

CHAPTER II - METHOD 19

Data 19
Cluster Analysis Procedures 20
Number of Clusters and Internal Validity 20
Cluster Structure Consistency 22
External Comparisons 23

CHAPTER III - RESULTS 25

Internal Validity 25
Consistency Analyses 26
External Comparisons 33

CHAPTER IV - DISCUSSION AND CONCLUSIONS 37

Internal Validity 37
Cluster Structure Consistency 38
External Comparisons 39
Limitations of the Research 41
Future Research 41

REFERENCES 45

APPENDIX A. MOS TITLES, APTITUDE AREAS AND CAREER MANAGEMENT FIELDS A1

B. WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA B1

C. AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA C1

D. K-MEANS CLUSTER ANALYSIS OF DOT DATA D1

E. WARD HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA E1

Page

APPENDIX F. AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA F1

G. WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA G1

H. AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA H1

I. K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA I1

J. WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA J1

K. AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA K1

L. K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA L1

LIST OF TABLES

Table 1. Enlisted Career Management Fields 12

2. Army Aptitude Areas 13

3. Enlisted DoD Occupational Codes 14

4. Number of Clusters Selected by Three Cluster Analysis Methods for DOT, Project A, and Synthetic Validity Data Bases 26

5. Rand Values Comparing DOT Cluster Structures Across Cluster Analysis Methods 27

6. Rand Values Comparing Project A Cluster Structures Across Cluster Analysis Methods 27

7. Rand Contingency Table Comparing DOT Ward and K-Means Solutions 28

8. Rand Contingency Table Comparing DOT Ward and Average Linkage Solutions 29

Page

Table 9. Rand Contingency Table Comparing DOT Average Linkage and K-Means Solutions 29

10. Rand Contingency Table Comparing Project a Ward and Average Linkage Solutions............................ 30

11. Rand Values Comparing Sample One and Sample Two Cluster Solutions in DOT Double Cross-Validation Analysis.. 33

12. Rand Values Comparing DOT Empirical Cluster Solutions With Nine Army Aptitude Area Job Families............ 33

13. Rand Values Comparing Project A Empirical Luster Solutions With Nine Army Aptitude Area Job Families.. 34

14. Rand Contingency Table Comparing DOT Ward and Army Aptitude Area Job Families........................... 34

15. Rand Values Comparing DOT and Project A Cluster Structures.. 35

LIST OF FIGURES

Figure 1. Number of clusters in DOT population and cross samples one and two by cluster analysis method............... 32

CHAPTER 1

INTRODUCTION

This project had two objectives. The first was to study the major issues and trends in job clustering research. The second was to develop a method for evaluating job family structures that would account for the numerous clustering techniques available, and the variety of purposes for which job families are used in personnel psychology.

This report provides the results of the study. Chapter One presents a review of the job family literature, and a description of Army job family structures. Chapter Two describes the proposed cluster evaluation method. Chapters Three and Four present the results and discussion, respectively, of the application of this method to four existing Army data bases of job description information.

Review of Job Family Research

Pearlman published a comprehensive review of job family research in 1980 that made several contributions to the literature. First, he presented a hierarchical classification of the content of work, placing job families at the most global level of analysis. His taxonomy begins at the molecular level with tasks and moves to more general positions, jobs, occupations and job families. He defined job family as a group of two or more jobs that are similar in either worker characteristics or job performance requirements.

Second, Pearlman (1980) presented a broad analytical framework for directing future research by identifying three basic issues that must be addressed in job family (or any classification), research: 1) the purposes for which job families are used, 2) the content of the groups, and 3) the clustering procedures. Third, his review of studies in public and private sector organizations demonstrated that job families support most of the basic personnel management procedures: vocational guidance, recruiting and selection, military personnel classification, training curriculum design, job evaluation, performance appraisal, career planning and development, and personnel research.

Harvey (1986) conducted a second review of the job family literature. He emphasized the procedures for forming job clusters, while Pearlman (1980) focused more on the content and purposes of job families. Harvey observed that the grouping of jobs into families is essentially a data reduction strategy that can improve our understanding of the structure of work by reducing the complexity of the problem. Further, the development of job families is a tool for simplifying the design of personnel procedures. Harvey (1991) pointed out that the important question in job classification research is not whether jobs are similar or different in absolute terms, but on which dimensions they are similar and on which they differ.

Both Pearlman (1980) and Harvey (1986), reviewing much the same literature, found that the number and composition of job families varied according to the choice of job analysis data, the clustering method, and the personnel procedure. Pearlman concluded that the purpose for which job families are formed must guide the choice of the job descriptor and the clustering procedure. Harvey (1991) stated that the goal of the development

process should be to determine which jobs can be treated interchangeably for a given purpose.

Sackett (1988) added to the discussion by stressing the need for validating job family structures against criteria related to specific personnel procedures (Pearlman, 1980, made this point earlier). He described the task of the personnel psychologist this way:

> The questions left unanswered are: Given that an appropriate job descriptor has been chosen, how large a difference between jobs on the chosen descriptor is needed to have a significant impact on the criterion of interest? In a selection setting, how different do jobs have to be before validity coefficients are affected? In a training situation, how different do jobs have to be before separate training programs are required? In a performance appraisal situation, how different do jobs have to be before separate performance ratings forms need to be constructed? Thus job clustering can only be meaningfully done with reference to an external criterion. (pp. 51-52)

Job family research has received less attention than other areas of personnel psychology. For example, only two published studies were found since Harvey's 1986 review. Despite Pearlman's (1980) attempt to place job family research on a firm theoretical footing, most studies have focused on examining methodological issues, e.g., the effect of alternative job descriptors on cluster structure. Zimmerman, Jacobs and Farr (1982) observed that the development of methods for constructing reliable, valid cluster structures is a necessary component of theory building. However, few attempts have been made to integrate research findings (Pearlman, 1980 and Sackett, 1988, are exceptions), as the basis for formulating a conceptual framework for job family research. Pearlman's statement that, "A central and long-standing problem in the study of human work performance has been the lack of a comprehensive system for classifying and interrelating performance-related variables" (p.1), still holds today.

The objective of the present research review is to describe alternative strategies for developing and validating job families, with emphasis on the major problems encountered, trends in the findings, and cumulative knowledge gained that would be useful for guiding future research. The job family literature was divided into three broad streams of research. The first group of studies describe the formation of job families for alternative personnel processes, e.g., selection and performance appraisal. The second group investigated the effect of varying the type of job descriptor on the number and composition of job families. The last group of studies involved evaluation of alternative clustering methods.

Clustering Jobs to Support Alternative Personnel Procedures

Much of the research within the last 15 to 20 years has been conducted in the private sector, and has responded to the need for legally defensible selection procedures. Selection involves grouping jobs for one of two objectives. The first is to group similar jobs together to obtain an adequate sample size for test validation. The second objective is to evaluate whether test validity generalizes to a new job not included in the study, or to similar jobs within or across organizations (Pearlman, 1980).

Cornelius, Schmidt & Carron (1984) stated that the formation of job families is one of the most important concerns in validity generalization research. Further, they noted that there is a basic discrepancy between scientific and legal guidelines for the development of job clusters. The Uniform Guidelines on Employee Selection Procedures (1980) recommend that an elaborate, behaviorally-oriented job analysis procedure be conducted to substantiate job similarities. In contrast, Cornelius et al. showed that global ratings of job similarity may be adequate to group jobs for validity generalization in many situations. Further, other studies (e.g., Sackett, Cornelius & Carron, 1981, and Stutzman, 1983), indicate that emphasis on work behaviors may not be appropriate for all types of selection problems.

Mobley and Ramsay (1973) conducted one of the early selection/validity generalization studies using the Ward hierarchical cluster analysis (HCA) procedure to cluster jobs in two chemical plants. HCA refers to a group of clustering algorithms that sequentially agglomerate objects (in our case jobs), into nested clusters based on some criterion of similarity. The process begins with every job as a cluster. The clusters are successively grouped together to form larger, more heterogeneous clusters, until only one cluster remains. The output of HCA procedures is a sequence of nested, nonoverlapping groups.

A large number of clustering criteria have been developed, but no one approach has been found to be best for all situations (Jain & Dubes, 1988). Each algorithm imposes a specific definition of a cluster on the data. The Ward procedure forms clusters that minimize the within-group sum of squares, or error sum of squares (Aldenderfer & Blashfield, 1984). Another hierarchical procedure, called the average linkage algorithm, forms groups by comparing two clusters in terms of the average distance between all pairs of objects.

Mobley and Ramsay (1973) clustered all jobs (about 60 per plant) across all levels in the two plants. The job descriptors were 15 worker attributes and 5 job characteristics. They obtained 4 and 6 clusters, respectively, in the two plants. The clusters varied in complexity and level of responsibility, and were consistent across the plants. The researchers concluded that the Ward procedure was useful for grouping jobs to increase sample size and for validity generalization.

Taylor and Colbert (Colbert & Taylor, 1978; Taylor, 1978; Taylor & Colbert, 1978) conducted a three part validity generalization study for a national insurance company. Job analysis information was obtained with the worker-oriented Position Analysis Questionnaire (PAQ) and a company-specific PAQ. The Ward minimum variance procedure, combined with a reassignment process, was used for clustering. One of the major weaknesses of hierarchical clustering algorithms is that once an object is assigned to a cluster at some level in the hierarchy, it cannot move to another cluster. As objects are added at higher levels, the cluster centroids will change and some of the objects may be closer to other groups. The reassignment process alleviates this potential weakness.

Two samples of jobs were obtained. The first included 76 jobs in 23 locations. Six fairly global job families, based on the PAQ dimensions, were obtained. The reliability of the cluster structure was assessed by

reclustering the jobs using PAQ ratings from a second sample of SMEs. Eighty-nine percent of the jobs were placed in the same job clusters.

The second sample consisted of 325 jobs at all levels below vice president in the 23 regional offices. Company-specific PAQ-type job descriptors were introduced in an effort to obtain more homogeneous and organizationally meaningful job families. The second study obtained 13 homogeneous clusters. However, 14% of the jobs were isolates or could be clustered into more than one job family. Further, the consistency of the groups was reduced to a 58% hit rate.

The final component of the research involved the development of selection batteries for 3 of the 13 clerical job families from the second sample. Colbert and Taylor (1978) found substantial support for generalizing validities within the job families. First, the cross-sample predictive validities within job families were statistically and practically significant, averaging around $R = .30$. Second, different combinations of predictors were valid for different job families. Third, the job families were found to moderate validity, i.e., the predictive validity coefficients of the test batteries were larger within job clusters than across.

Alley, Treat, and Black (1988) tested the Ward HCA procedure for grouping Air Force jobs into families used for military personnel classification. They clustered the predicted performance scores of Air Force enlisted personnel in over two hundred specialties. The performance estimates were obtained by regressing final training grades against the 10 tests of the Armed Services Vocational Aptitude Battery (ASVAB).

Six job families were obtained. Four of the clusters matched the existing Air Force families labeled Mechanical, Administrative, General and Electronics. One of the additional families was a generalist cluster that required high scores on all 10 tests. The second new grouping was a heterogeneous cluster of jobs for which the ASVAB was not a good measure of training performance. Although they did not validate the clusters against a measure of classification efficiency, Alley et al. (1988) found that the least squares estimates of training performance for the new configuration of 6 job families provided significantly greater predictive validity than a single common equation.

Ballentine, Cunningham, and Wimpee (1992) evaluated the usefulness of a multipurpose job analysis instrument, the General Work Inventory (GWI), for forming career fields. The GWI includes broad worker- and job-oriented items, and was designed as an alternative, and supplement, to task information obtained from the Comprehensive Occupational Data Analysis Programs (CODAP) (Christal & Weissmuller, 1988). Twenty-one clusters, which accounted for 90 percent of the jobs, were obtained with the Ward HCA procedure.

The quality of the job clusters was evaluated by measuring the stability of the solution through cross-validation. Further, external validity was assessed by comparing the new clusters to the Air Force's existing career fields. The authors (Ballentine et al., 1992) found meaningful, stable clusters, but low overlap with the existing career fields. The last finding calls into question the use of a multipurpose job analysis inventory for developing a job family structure for a specific purpose, e.g., career guidance.

Several other recent studies have examined job families in military settings. Research using Army data is described in the section on Army job families. A study conducted for the Navy is discussed in the next section.

A group of studies published in the late 1970's and early 1980's attempted to develop what Harvey (1986) refers to as inferential job clustering procedures. The objective of these methods, which include analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA), was to provide a basis for evaluating the statistical significance of differences in job clusters. The inferential clustering procedures were developed to overcome two limitations of hierarchical and partitional clustering algorithms, which do not have built-in methods for statistical significance testing.

First, as noted by Jain and Dubes (1988), "Clustering methods have the nasty habit of creating clusters in data even when no natural clusters exist, so hierarchies and [partitional] clusterings must be viewed with extreme suspicion" (p. 75). Second, these algorithms determine the relative, not absolute, similarity, among objects (Sackett, 1988). These limitations are particularly serious when attempting to develop job families for defensible testing and performance appraisal systems. Inferential clustering procedures received a great deal of attention as potential solutions to the problems, but have faded from the literature due to their serious weaknesses.

Arvey and Mossholder (1977) proposed using two-way ANOVA with repeated measures to group jobs into job families. The levels of the first factor are the job descriptors, the levels of the second factor are the jobs. Multiple raters nested within jobs form the third factor. The jobs main effect indicates whether there are mean differences in the jobs collapsed across job descriptors. If the jobs main effect is not significant, then the jobs form a single job family. If it is significant, then post hoc comparisons are used to identify job families. The jobs x descriptors interaction provides information on the pattern of job analysis dimensions that accounts for the differences in job means.

Lissitz, Mendoza, Huberty and Marlos (1979) criticized the ANOVA approach saying that it is unreasonable to treat job descriptors as levels of a single independent variable, instead of as separate dimensions. They proposed a MANOVA design in which the job analysis items are individual dependent variables. Arvey, Maxwell and Mossholder (1979), and Lee and Mendoza (1981), refined the discussion by describing the strengths and limitations of both ANOVA and MANOVA, and identifying the conditions under which each procedure would be most appropriate. Arvey, Maxwell, Gutenberg, and Camp (1981) later conducted a Monte Carlo study of the ANOVA approach and found that omega-squared estimates for the jobs x descriptors interaction were more useful than statistical significance tests for detecting differences in jobs.

Despite the original appeal of the inferential procedures for job clustering, numerous criticisms have been raised about their efficacy for this purpose (Hanser, Mendel, & Wolins, 1979; Harvey, 1986; McIntyre & Farr, 1979). The most important limitation of the ANOVA approach was mentioned above, i.e., the assumption that the job analysis dimensions are levels of a single variable. Another problem with both techniques is that the capacity for finding significant differences is partly a function of the power of the test.

In other words, with large sample sizes even small differences will be significant, while with small samples true differences may go undetected. Finally, the procedures are highly impractical with more than a few jobs, because the clustering is done by the researcher through visual inspection of the ANOVA tables for n(n-1)/2 pairs of jobs (where n is the number of jobs). This process is handled quantitatively by hierarchical and partitional clustering algorithms.

Cornelius, Carron and Collins (1979) developed a different approach to estimating the statistical differences in clusters by following a hierarchical cluster analysis with discriminant analysis. Although this approach is intuitively appealing, it has also been criticized. Numerical taxonomists state that conducting a discriminant analysis, or ANOVA or MANOVA procedure, based on groups obtained from an empirical cluster analysis, is inappropriate, because it will almost always produce significant results (Aldenderfer & Blashfield, 1984; Milligan & Cooper, 1987). Milligan and Cooper (1987) explained that discriminant analysis and MANOVA are biased toward finding significant results because the groups are not defined a priori. Additionally, the statistical tests used to evaluate the significance of the results are biased, because the variables in the model are the same ones used to define the clusters. However, discriminant analysis can be used in job family research as a method for placing new jobs in previously existing groups (Cornelius et al., 1984).

The Relationship of Job Descriptor to Job Family Structure

The second group of five studies examined the effect of job analysis information on cluster structure. Cornelius et al. (1979) investigated the effect of three alternative types of job descriptors (tasks, behaviors and job-related abilities), on the number and composition of job families. The purpose of the study was to determine whether a single selection test could be used with seven functionally different foreman jobs in a chemical processing plant.

The Ward HCA procedure, combined with discriminant analysis, was used to form groups. Three distinct cluster structures were obtained. The task-oriented items produced the greatest number of clusters, between 3 and 5; the behaviorally-oriented items (from the PAQ) produced 1 global cluster; and the ability items produced 3 clusters.

The authors (Cornelius et al., 1979) noted that the task cluster results demonstrated one of the major limitations of HCA. Examination of the plot of within-cluster sum of squares by number of clusters showed that two cluster solutions (3 and 5 groups), were equally homogeneous. Consequently, the choice of a single structure could not be made on the basis of the cluster results alone. The final selection of 5 clusters entailed analysis of the composition of the clusters of each solution, thereby introducing subjective judgement into the clustering procedure.

Cornelius et al. (1979) concluded that the job descriptor plays a major role in determining the number and composition of job families, and that the match between the type of data and the personnel procedure must be a major consideration in designing studies. They cited, for example, two types of selection problems that require different descriptors. When selection is for jobs that require previous education and training, and the applicant is

expected to perform the activities immediately, task-oriented descriptors would be most appropriate for developing selection instruments and job families. When selection is followed by training in job-related knowledge and skills, as in the Armed Services, then jobs should be grouped by broad cognitive, perceptual and psychomotor abilities (Sackett, 1988).

Reynolds, Barnes, Harris and Harris (1992) also investigated clusters formed by tasks, behaviors and abilities, but for a sample of 75 Navy jobs that varied in function and rank. The Ward minimum variance technique was used in combination with the K-means partitional clustering procedure. Iterative or partitional clustering algorithms generate a single partition of the data into *K* clusters. They were developed to improve upon hierarchical methods that do not permit reassignment. The typical procedure begins with a random partition of objects into clusters. The centroids of the clusters are computed, and the objects are reassigned to the cluster with the nearest centroid. Then the centroids are recalculated and the objects shifted once again. This process continues until no objects change clusters. In the K-means procedure, as in the Ward algorithm, objects are assigned to clusters to minimize the within-cluster sum of squares.

Like Cornelius et al. (1979), Reynolds et al. (1992) found different cluster structures by type of descriptor, but the pattern of results was different. Where Cornelius et al. obtained only 1 cluster for behavioral descriptors, Reynolds et al. obtained 9, the largest number in the study. Six clusters were found for both task and ability information, but the composition of the families differed. The task clusters were grouped along functional dimensions: electronics, construction, administrative, weapons, machinery, and communications/ cryptology. The ability-based families were more abstract: perceptual/cognitive, communications/dexterity, cognitive/dexterity, communications, average physical, and physical.

In the third study addressing the relationship of descriptors to cluster structure, Cornelius, Hakel and Sackett (1979) used three-mode factor analysis to develop job families to support the design of a performance appraisal system for the Coast Guard. The objective was to identify the number of different appraisal forms needed for all enlisted jobs across functions and levels. Before the study a single form was used for all jobs and levels. The modes of the analysis were: PAQ behaviorally-oriented items, job titles, and level of responsibility (rank) within the organization.

The approach was to compute a rotated factor solution for each mode, or job descriptor. A core matrix was then obtained that related the factors of the three modes. The authors (Cornelius et al., 1979) described the analytical model as a three-dimensional cube. The first side consisted of 18 levels for the 18 jobs sampled. The second side had 5 levels that accounted for the 5 ranks included in the study. The third side of the cube had 153 levels representing the 153 PAQ items. Each cell in the cube consisted of a mean relative time spent value that was a unique combination of job, rank and PAQ elements. The three-mode factor analysis produced a 5- by 2- by 7-factor solution. In other words, there were 5 global job families, which were subdivided into two (high and low) ranks, and which differed in their patterns on 7 PAQ-based job analysis factors.

Cornelius et al. (1979) considered this approach to be superior to HCA for developing a performance appraisal system for several reasons. First, all

three types of data were used to identify and describe the families. Second, the factor loadings were easier to interpret than cluster profiles from HCA. However, interpretation of the factor loading matrix would become difficult if there were a great deal of overlap in the families. Third, the pattern of factor loadings across the modes provided a rich source of data for developing the appraisal instruments.

In the final two studies that examined different types of job descriptors, Sackett et al. (1981) and Cornelius et al. (1984) investigated whether holistic ratings of job similarity would produce the same job families as a molecular approach. Both studies found high consistency between global and specific ratings. In the Cornelius et al. study, task- and worker-oriented activities were compared to ability elements and a holistic classification of jobs. More stable results were obtained with the activities-based clusters than with the ability-based job families, which showed substantial overlap. The holistic families were similar to the activity clusters, and the authors concluded that the holistic rating procedure was equivalent to the more resource intensive activity profiles. However, the results for the holistic clusters are questionable, because the ratings were made after completing the full job analysis questionnaire. We would expect that the raters' judgements about the molecular content of the jobs would influence their final global ratings.

Comparative Analyses of Job Clustering Procedures

Another line of research is the comparative analysis of alternative clustering methods. This research has focused mainly on hierarchical and partitional cluster analysis. Two studies were found that compared alternative algorithms, and attempted to make generalizations about the best procedures for different job clustering situations. Monte Carlo procedures based on synthetic data were used to build in experimental controls of the independent and dependent variables, e.g., number of jobs, variables and clusters, types of error and outliers, and the shapes and relative sizes of clusters. Consequently, the results provide only limited guidance for applied research.

Garwood, Anderson and Greengart (1991) compared three HCA algorithms (Ward minimum variance, average linkage and single linkage), and two similarity measures (the correlation coefficient and Euclidean distance). The purpose of the study was to evaluate the effects of alternative job family structures on the accuracy of different algorithms in recovering the true structure. Research in numerical taxonomy and cluster analysis has shown that the shapes, relative sizes, and amount of overlap of clusters differentially affects the ability of clustering algorithms to identify the true clusters (Blashfield, 1976; Milligan, 1981a, 1981b; Milligan & Cooper, 1987; Milligan, Soon & Sokol, 1983; Sarle, 1983). Garwood et al. studied this problem specifically as it applies to jobs.

The researchers (Garwood et al., 1991) used a Monte Carlo mixture model procedure to generate four different types of cluster structures. In cluster analysis a mixture model refers to the generation of artificial data that can be viewed as samples from a mixture of different multivariate clusters of objects (Milligan & Cooper, 1987). The cluster structures generated by Garwood et al. (1991) varied in the relative sizes of the groups, the shapes of the distributions, and the levels of mean differences in profile scores.

Four types of job family structure were simulated: 1) increasing responsibility, as in career families, 2) specialists among generalists, as in a research group that includes computer programmers, 3) different functional areas, and 4) hierarchical functions with overlapping responsibilities across levels, e.g., nurses.

The results showed that average linkage was slightly better than the Ward algorithm for all types of cluster structures; while single linkage was a distant third. Further, average linkage performed well when the data contained outliers. The correlation coefficient produced the best recovery when clusters differed by shape, and when the data contained outliers. Distance was superior when clusters differed by level of job analysis profile. (Two job families would differ by level, for example, if the jobs in one family were more difficult than in the other.) The efficiency of both the Ward and average linkage procedures was reduced by differences in cluster size.

The findings have important implications for matching the clustering procedure and similarity measure with the personnel procedure for which the job families are created. For example, a study to design a career development program would probably sample jobs with increasing responsibility within one functional area. In this case the average linkage procedure, with a distance measure, would be expected to provide the best recovery. If job families were being constructed to support the design of performance appraisal instruments, the sample might include jobs from all functional areas at the same level in an organization. The correlation coefficient would be the appropriate proximity measure in this situation.

Zimmerman et al. (1982) compared the Ward HCA algorithm to a combination of Ward plus the K-means iterative procedure, and a density search technique. Density search algorithms are designed to identify natural clusters of any shape by analyzing areas of greater and lesser density. Theoretically, they should be superior to Ward and K-means clustering when the relative sizes of groups differ. Using a mixture model of multivariate normally distributed clusters, the authors examined the recovery of alternative procedures when the number of jobs, clusters and job descriptors, and the relative sizes of the clusters, were varied.

Contrary to expectations, the Ward and Ward/K-means procedures were generally about equal in recovering the true cluster structures, while the density search technique was significantly worse. The Ward/K-means procedure was expected to be superior to the Ward procedure alone, especially with a large number of jobs, because of the reassignment feature. Increasing the number of job descriptors and the number of clusters improved the first two procedures, but not the latter. Increasing the number of jobs had a positive effect on only the K-means approach. The authors also examined the effect of overlapping clusters on the three methods. Only the Ward procedure was reduced in accuracy with overlapping clusters.

The Garwood et al. (1991) and Zimmerman et al. (1982) studies fit within a larger body of cluster analysis research that spans many fields in the social and physical sciences. Their results are partially supported by comparative analyses outside of personnel psychology. However, the larger body of cluster analysis research contains a great many inconsistencies and is

difficult to summarize. This is partially due to the lack of consistency in methods, designs and variables under investigation.

A small indication of the conflicting results that characterize the cluster analysis literature can be seen by comparing the above mentioned studies to a comparative analysis by Scheibler and Schneider (1985). They investigated nine hierarchical and four nonhierarchical clustering procedures using synthetic data. Like Garwood et al. (1991), they found that the correlation coefficient was generally more effective than a distance measure in recovering cluster structure. They speculated that this could be because a correlation, which is insensitive to mean differences in the variables, may be more robust in the presence of outliers.

In contrast to Garwood et al. (1991) and Zimmerman et al. (1982), Scheibler and Schneider (1985) found that the Ward algorithm, used with a distance measure, outperformed both a combined Ward/K-means approach and average linkage. A second conflict was that the recovery of average linkage was poor with a distance measure. Finally, Scheibler and Schneider also compared the Ward and average linkage procedures when clusters overlapped. Unlike Zimmerman et al, they found that the Ward algorithm was more accurate than average linkage.

Conclusion

The issue that received the most attention in recent job family research was the effect of the job descriptor on cluster structure. Several conclusions were drawn from the studies that provide guidance for future research. First, Sackett (1988) stated that the level of specificity of the job descriptors should match the level of specificity of the personnel procedure for which jobs are clustered. For example, global ratings of the similarity among jobs would not be appropriate for grouping jobs into families for training. Second, broad and specific descriptors have both strengths and weaknesses, and should be considered complementary, not competitive. More technically and behaviorally abstract descriptors (e.g., the PAQ, abilities), produce more global and overlapping job families (Harvey, 1986, 1991). Specific descriptors (e.g., tasks), tend to produce a larger number of more narrowly defined clusters, but this depends on the construction of the items. Third, using a large number of job analysis dimensions appears to improve the quality of the cluster solution (Zimmerman et al., 1982).

Despite recognition of the importance of the choice of job descriptor for cluster structure, more research is needed to determine specifically which descriptors are appropriate for different personnel processes. Beyond the general guidelines mentioned above, the choice at present depends almost entirely on the judgement and experience of the researcher.

A large number of clustering methods have been found useful for forming job families,[1] but studies based on actual jobs and descriptors have tended to investigate only one clustering procedure for a specific purpose. Comparative analyses of clustering procedures in applied settings are needed. Garwood et al. (1991) and Zimmerman et al. (1982) addressed this issue in

[1]See Harvey (1986) and Schoenfeldt (1985) for more extensive descriptions of clustering procedures than were included here.

well-controlled Monte Carlo studies. However, the generalizability of their results to applied data is unknown.

Finally, most of the applied job family studies attempted to validate cluster structures against an external criterion. A small number of studies evaluated the stability of clusters. However, no systematic method for evaluating the quality of cluster structures was found. This was true in both the applied and Monte Carlo research domains.

An applied cluster evaluation method, described below in Chapter Two, was developed in this study to address the gap in the literature. The proposed method was used to conduct a comparative analysis of hierarchical and partitional clustering procedures. The job family structures of four Army data bases containing actual jobs and alternative job descriptors were investigated. The method was assessed for its utility in constructing reliable, valid cluster structures in applied research settings.

Army Job Families

Army Career Management Fields

According to Army Regulation (AR) 611-201, Enlisted Career Management Fields and Military Occupational Specialties, the Commanding General, U.S. Total Army Personnel Command (CG, PERSCOM) develops and maintains the career management fields (CMF) and military occupational specialties (MOS) for five major purposes: recruitment, training, assignment, distribution, and professional development. The CG, PERSCOM is responsible for establishing "a methodology for review, analysis, and implementation of classification structure changes" and for maintaining "military career progression patterns" (p. 17, Headquarters, Department of the Army, 1992). PERSCOM operates under the guidance of the Deputy Chief of Staff for Personnel (DCSPER), who establishes occupational classification structure policy (Headquarters, Department of the Army, 1992).

The objectives of the CMF structure are to: provide logical patterns for progression to higher level jobs, provide standard grade-skill level relationships, consolidate MOS at higher grade levels, and provide for self-sustainment through entry-level recruiting or lateral entry from other CMF where appropriate. Currently, there are 35 CMF as shown in Table 1.

The CMF and MOS are part of the Army's Enlisted Personnel Management System (EPMS). Rosse, Borman, Campbell and Osborn (1983) described the CMF from an historical perspective. They stated that the CMF system was developed as part of the branch system, which contained 14 groupings, or branches, of soldiers that represented a specific service of the Army, e.g., Infantry or Finance. The CMF are designed to contain homogeneous sets of MOS which share common career paths, but practical considerations also determine their composition. The MOS within a particular CMF are related to ensure that soldiers have the knowledge and abilities for training and assignment to other specialties within the CMF.

Rosse et al. (1983) noted three practical considerations, in addition to similarity of ability patterns, that help to determine the composition of the CMF. First, the CMF support the "feeder-capper" system. This system is designed to provide a career path for all MOS from E-1 to E-9. MOS lacking

Table 1. Enlisted Career Management Fields

- Administration
- Recruiting and Reenlistment
- Public Affairs
- Bands
- Aircraft Maintenance
- Aviation Operations
- Civil Affairs
- Electronic Maintenance and Calibration
- General Engineering
- Chemical
- Topographic Engineering
- Medical
- Mechanical Maintenance
- Electronic Warfare/Intercept Systems Technology
- Military Intelligence
- Signals Intelligence/Electronic Warfare Operations
- Military Police
- Infantry
- Combat Engineering
- Field Artillery
- Air Defense Artillery
- Special Forces
- Armor
- Air Defense System Maintenance
- Land Combat and Air Defense System Direct and General Support Maintenance
- Psychological Operations
- Visual Information
- Signal Maintenance
- Signal Operations
- Record Information Operations
- Ammunition
- Supply and Services
- Petroleum and Water
- Food Services
- Transportation

Source: Headquarters, Department of the Army, 1992.

positions for higher grades are feeders and are assigned a capper MOS (with somewhat similar job content), which does have higher grades. The feeders and cappers are placed in the same CMF so that soldiers in the feeder MOS can advance to E-9 while remaining in the same CMF. Second, MOS that entail working with the same types of equipment are often placed in the same CMF. Third, some CMF are designed to represent a mission identity, e.g., Special Forces, that includes a heterogeneous group of MOS.

In summary, the Army's CMF system is a personnel management tool used for developing, counseling, and managing enlisted personnel. Although homogeneity of job content is a primary concern in configuring CMF, practical considerations tend to increase their heterogeneity.

Army Aptitude Areas

The Army Aptitude Area (AA) system was developed in 1949 when the Army Classification Battery (ACB) formed the basis of selection and classification testing (Johnson, Zeidner & Leaman, 1992). At that time there were 10 AA, or job families. In 1956 the number of AA was reduced to 7.

In 1972 Maier and Fuchs increased the number of AA to 9 as part of a selection and classification efficiency study to revise the ACB (see Table 2 below). The revision of the AA system was based on a partly empirical and partly rational procedure. Both subject matter expert (SME) judgements of the similarities in performance requirements, and the similarity of patterns of weights on cognitive and interest measures, were utilized in forming the families. In addition, practical considerations, e.g., adhering to the official Army organizational structure, were included wherever possible.

Although the ASVAB has replaced the ACB as the Army's selection and classification battery, the AA established in 1972 are still in use (Maier & Grafton, 1981). Each AA is associated with a unique composite of 3 or 4 unit-weighted ASVAB subtests. The cutoff scores on the composites determine a

Table 2. Army Aptitude Areas

• Combat	Infantry, Armor, Combat Engineer
• Field Artillery	Field Cannon and Rocket Artillery
• Electronics Repair	Missiles Repair, Air Defense Repair, Tactical Electronic Repair, Fixed Plant Communications Repair
• Operators and Food	Missiles Crewman, Air Defense Crewman, Driver, Food Services
• Surveillance and Communications	Target Acquisition and Combat Surveillance, Communication Operations
• Mechanical Maintenance	Mechanical and Air Maintenance, Rails
• General Maintenance	Construction and Utilities, Chemical, Marine, Petroleum
• Clerical	Administrative, Finance, Supply
• Skilled Technical	Medical, Military Policeman, Intelligence, Data Processing, Air Control, Topography and Printing, Information and Audio Visual

Source: Maier & Fuchs, 1982.

recruit's eligibility for entry into specific MOS. Johnson et al. (1992) recently evaluated the effectiveness of the AA composites and job families for classification. They found that increasing the number of groups from 6 to 9, 12, 16 and 23, and altering their composition to increase the homogeneity of the jobs in terms of ability profiles, produced statistically and practically significant increments in classification efficiency. In addition, Johnson et al. compared the AA to the more numerous CMF in terms of classification efficiency, and also found a large increase in classification efficiency when the CMF were used for job assignment.

The Johnson et al. (1992) study has been criticized because the purely empirical configuration of their classification-efficient job families combines jobs with highly divergent content into a single cluster (Laurence & Hoffman, nd). Consequently, recruiting and classification with the Johnson et al. job clusters would be difficult from a marketing perspective. For example, one family included M48-M60 Armor Crewmember, Food Service Specialist and Military Police. Recruiters would have a tough time persuading applicants to enter some of these analytically-determined families. In general, as more families were included in the Johnson et al. (1992) configurations, they became more functionally similar and operationally meaningful.

The major benefit of the recent classification work is that it indicates that the efficiency of the Army classification system could be improved by increasing the number of job families, and modifying their composition to maximize classification efficiency. But practical concerns, e.g., marketability to recruits, must also be considered in the development process.

Enlisted DoD Occupational Codes

The Department of Defense (DoD) has constructed a classification of military occupations that reflects the job family structures of the four Services (Office of the Assistant Secretary of Defense [Force Management and Personnel], 1987). This classification is a hierarchical arrangement consisting of three levels. At the most global level there are 10 occupational groups with corresponding one-digit codes as shown below in Table 3.

Table 3. Enlisted DoD Occupational Codes

- Infantry, Gun Crews, & Seamanship Specialists
- Electronic Equipment Repairers
- Communications & Intelligence Specialists
- Health Care Specialists
- Other Technical & Allied Specialists
- Functional Support & Administration
- Electrical/Mechanical Equipment Repairers
- Craftsmen
- Service & Supply Handlers
- Non-Occupational

Source: Office of the Assistant Secretary of Defense, [Force Management and Personnel], 1987).

The 10 global clusters are further subdivided into 68 more narrowly-defined clusters with two-digit codes. The most specific level of the taxonomy includes 160 subgroups with three-digit codes. Allocation to occupational codes is based on "careful analysis of the duties of each specialty" across the Services (Office of the Assistant Secretary of Defense [Force Management and Personnel], 1987).

The DoD occupational groups are useful for identifying jobs across Services that fall into the same job families. They also have been used in selection and classification research to reduce the complexity of the analysis (Harris, McCloy, Dempsey, Roth, Sackett, & Hedges, 1991).

Job Families in Research

A number of different job family structures have been constructed as part of selection and classification research projects. Selection of the sample of 21 MOS for the Army's comprehensive Selection and Classification Project (Project A), was based on a combination of empirical cluster analysis and practical criteria (Hoffman, 1987; Rosse, Borman, Campbell & Osborn, 1983). The multi-step process combined the efforts of project scientists, Army officers and expert review panels. Campbell (1990) stated that the criteria for determining which MOS became part of Project A were:

1) oversampling of high density MOS with large numbers of women and minorities, 2) oversampling from the population of the most critical MOS, and 3) inclusion of jobs that represented major job families clustered by judged similarity of task content.

The job clustering component of the sample selection process was conducted in two phases (Hoffman, 1987; Rosse et al., 1983). In Phase One, a sample of 111 MOS was selected from the larger population of jobs, because pilot testing indicated that rating the total population of more than 200 entry-level MOS would be an unmanageable task for SMEs (Rosse, 1983). Based on the research of Cornelius et al. (1984), Rosse et al. (1983) used a global rating task instead of conducting a more time-consuming and expensive task analysis. Military and civilian SME sorted MOS into groups based on job titles and the similarity of performance requirements taken from AR 611-201. The interrater reliability was .94.

The SME sorting task was used to form a similarity matrix in a two-stage process. The first matrix contained elements that were the proportion of raters who grouped a pair of jobs into the same family. The vectors of this matrix for all pairs of jobs were correlated to form the final similarity matrix. Job families were formed by evaluating orthogonal and oblique principal factor solutions of the correlation matrix.

A 23-cluster solution was derived from analyzing patterns of factor loadings of MOS across 15 factors. This solution was compared to the CMF using Cohen's (1960) kappa statistic. Kappa provides a measure of agreement between two matrices, with a correction for chance matches. The kappa value of .55 indicated moderate agreement between the two job family structures. The 23-cluster solution provided input into the selection of 19 MOS for the Project A concurrent validity study. The authors observed that a potential weakness in the clustering methodology was that the raters might rely more on the MOS titles to form clusters than on the job descriptions.

The Phase Two cluster analysis was designed to select two more MOS for subsequent Project A research (Hoffman, 1987). This time raters were asked to sort all 268 entry-level MOS, using the original 23 clusters to start. A latent partition analysis, like the one conducted in Phase One, was performed along with the average linkage HCA procedure. The second study also produced 23 clusters.

Peterson, Owens-Kurtz and Rosse (1991) later clustered the 21 Project A jobs as part of the Army synthetic validity project. The MOS were clustered on the basis of a 96-item task questionnaire for which ratings of importance for core-technical proficiency (CTI) and the overall job (OJI) were obtained. The Ward minimum variance procedure was used, with the correlations of task profiles forming the similarity matrix. Two different cluster solutions of 4 and 3 groups, respectively, were obtained. The CTI clusters were labeled: electronics, administration/support, combat and mechanical/construction. The OJI clusters were: electronics/repair, administration/support and combat.

Research Questions

The objective of this study was to develop and test a method for evaluating the quality of job family structures, and for comparing alternative clusters of jobs developed for different purposes (e.g., personnel

classification or career development), and based on different types of descriptors (e.g., tasks or aptitudes). Sokal (1988) identified validation and comparison of cluster structures as among the unsolved problems in numerical taxonomy. Relatively little work has been done on developing cluster evaluation methods, because most of the research has focused on the development of new clustering algorithms (Jain & Dubes, 1988).

One exception is the work of Milligan and his colleagues (Milligan, 1979, 1980, 1981a, 1981b, 1985; Milligan & Cooper, 1985, 1987; Milligan & Isaac, 1980; Milligan & Schilling, 1985; Milligan & Sokal, 1980; Milligan, Soon & Sokol, 1983). These researchers conducted a series of Monte Carlo studies using synthetic data, where the true cluster structure is known, to investigate alternative indexes for evaluating the internal and external validity of cluster structures. Our approach was taken largely from Milligan's work. The major difference is that the evaluation method developed in this study was designed to assess the quality of cluster structures based on actual data where the true cluster structure cannot be known. Not only will an applied cluster evaluation method make an important contribution to job family and cluster analysis research, it will be of immediate practical application to the Army, which uses job families as personnel management tools to make operational decisions.

The core research question underlying the method was how to reconcile the differences in the number and composition of job families found in the literature and within the Army. As discussed above, the Army has two formal job family structures (CMF and AA), designed to group jobs according to similarity in performance requirements for two different purposes: career progression and personnel classification. Several studies have constructed alternative configurations of jobs, e.g., Rosse et al. (1983), Hoffman (1987), Peterson et al. (1990), Johnson et al. (1992), and Statman (1993). At present there is no cluster evaluation method that can provide quantitative information about the differences in these job family structures. A second stimulus for investigating methods of assessing the composition of Army job families comes from the changes in the Army's mission, structure, and the design of jobs, due to the end of the cold war and the alteration of the U.S. role in world politics.

Zimmerman et al. (1982) observed that any method for assessing the quality of a job family structure must include measurement of stability or consistency, and of the accuracy of the solution in uncovering the true structure of the jobs. The proposed evaluation method consists of a set of tools (analytic procedures and indexes), that can be used in both basic research and applied studies, and includes the following components:

- internal validation,
- consistency analysis,
- external comparisons, and
- validation against an external criterion.

Internal validation of the cluster structure of a set of objects, in our case jobs, involves identifying the number and composition of clusters that provides the best recovery of the underlying relationships among the objects in the distance matrix (Jain & Dubes, 1988; Milligan, 1981a). This must be the first step in evaluating the quality of a cluster solution obtained by empirical or rational means.

The consistency of a cluster structure is its stability across alternative clustering methods or samples. Consistency analysis is a replication method of evaluating whether the observed cluster structure is the true cluster structure. The limitation of replication studies in applied data sets is that a negative finding provides little or no information about the sources of inconsistency (Milligan & Cooper, 1987). For example, if little overlap in cluster structure is found across clustering methods, the discrepancies may be due to the algorithms, or the characteristics of the data, or both. Consistency analysis provides no internal mechanism for determining the contribution of either source of error.

External comparisons in Monte Carlo studies involve validation of the empirical clustering procedures against the true structure of the data created by the researcher (Milligan & Cooper, 1987). An external validation can be conducted in an applied setting if an external standard can be designated as the true cluster structure. However, the objective of applied job family research is to group jobs into families when the true structure is unknown. Consequently, external comparisons in job family research are limited to evaluating the overlap of new clusters with existing operational job families (e.g., the AA or CMF), or with other cluster structures based on different job descriptors. Although these analyses are not validations, they provide diagnostic information about the extent of change that could be expected by substituting the new clusters for preexisting ones, or about the similarities and differences in the definitions of job families based on different dimensions of work, e.g., tasks, behaviors and aptitudes.

The first three components of the proposed cluster evaluation method were tested using available job analysis data on Army MOS. The method's usefulness as an analytical tool was assessed in terms of the quality of the data it produced, and the relevance of the procedures for answering important applied research questions.

The last step in the method, validation of clusters against an external criterion, was beyond the scope of this study. However, it is necessary in applied research when the clusters will be used in decision-making, because the true structure of the data is unknown. The external criterion must be the effectiveness of the cluster structures for accomplishing some operational purpose, e.g., assigning applicants to jobs in a personnel classification system, helping soldiers chart their career paths, or providing a framework for personnel training and development programs.

CHAPTER II

METHOD

Data

The study was designed to use existing job analysis data from three previous studies rather than to collect new data. The first data base was compiled in the Joint Service Job Performance Measurement Project (Harris, McCloy, Dempsey, Roth, Sackett, & Hedges, 1991), which included entry-level military jobs across all four Services. Harris, McCloy, Dempsey, DiFazio and Hogan (1993) used the Army jobs in a subsequent study of alternative selection and classification models. Only the Army jobs were examined in the present study.

Job descriptors were obtained for 263 Army MOS using job analysis information from the Dictionary of Occupational Titles (DOT) (Harris et al., 1991). The DOT data base of occupational codes and job analysis ratings on 44 items was obtained for civilian jobs from the National Technical Information Service (U.S. Department of Labor, 1977). These jobs were matched to all entry-level Army jobs in existence in the mid-1980's using a military-civilian crosscode data base (Lancaster, 1984; Wright, 1984). Only about 20 MOS were excluded from the analysis because they could not be matched to civilian jobs. However, several groups of MOS, e.g., many electronics jobs, received identical descriptors because the civilian job structure was not as differentiated as the Army MOS.

The 44 DOT items cover worker functions (the DOT data, people, things scales), training time, cognitive aptitudes, temperaments, interests, physical demands and working conditions. Harris et al. (1993) reduced the items to four Army-specific orthogonal principal components, rotated to varimax simple structure. The principal components accounted for about 50 percent of the variance in the job descriptors and were labeled: 1) working with things, 2) complexity, 3) unpleasant working conditions, and 4) stressful working conditions. This DOT data base was considered to be the population of Army jobs at data collection in the late 1980's.

The second data base was the subject matter expert (SME) sorting data from the Project A cluster analysis studies (Hoffman, 1987; Rosse et al., 1983) described above. It contained 268 entry-level MOS, and also represented the Army's job population at the time of data collection. The proportion of SMEs who assigned pairs of jobs to the same cluster, based on a global description of performance requirements, formed a similarity matrix that was the basis for clustering. The DOT and Project A data bases had 227 overlapping jobs. The differences reflected changes in the Army's job structure across the two data collections.

The third data base was the sample of 21 Army MOS studied in the Army synthetic validity project (Wise, Peterson, Hoffman, Campbell, & Arabian, 1991). This sample was partially based on the cluster analysis results obtained by Rosse et al. (1983) and Hoffman (1987). The 21 jobs span 16 of the 23 clusters they found, and was selected to meet a number of scientific and practical criteria (Campbell, 1990). The job descriptors were 96 task statements rated in importance for either core technical proficiency (CTI) or

for the overall job (OJI). The development and analysis of the task questionnaire is described in Hoffman, Fotouhi, Campshure, & Chia (1991).

Cluster Analysis Procedures

Three clustering algorithms were compared: Ward's minimum variance hierarchical cluster analysis (HCA), average linkage HCA, and a modified K-means algorithm contained in the SAS computer program (SAS Institute Inc., 1990). The only difference between SAS's K-means procedure and MacQueen's (1967) original algorithm is that initial cluster centroids are determined in a manner that minimizes the number of iterations needed to obtain a final solution. The K-means procedure was used with a set of random seeds, or initial cluster centroids. Three clustering algorithms were chosen for investigation because previous comparative studies, based on Monte Carlo analyses, have found that each one provides accurate cluster recovery under different conditions (see, for example, Zimmerman et al., 1982, and Garwood et al., 1991).[2]

In all but the Project A analyses, squared Euclidean distance, d^2, was the proximity measure. Squared Euclidean distance measures the differences in the level and shape of job descriptor profiles, but level is more important in determining cluster structure. The correlation coefficient is another common measure of proximity in cluster analysis. It measures only similarity in profile shape. A correlation coefficient may have been more appropriate with the data in the present study, because all jobs are at entry-level, and probably differ more in profile shape than in level. However, d^2 was the single proximity measure that would allow us to use the SAS cubic clustering criterion (CCC) for the internal validation. Initial analyses were conducted with both measures. Comparison of the cluster solutions showed that the clusters were different, but that neither measure was consistently superior.

The similarity measure for the Project A data was the proportion of raters who placed pairs of jobs in the same cluster. This measure was subtracted from one to produce a dissimilarity measure, as required by the SAS algorithms.

Number of Clusters and Internal Validity

Indexes

Two indexes, CCC and Hubert's (Hubert & Arabie, 1988) Gamma, were used to identify the best cluster solution from among a large number of alternative classifications. CCC is a measure of the proportion of variance accounted for by the observed cluster solution, compared to the proportion of variance accounted for by clustering a uniform distribution based on a hyperbox (Milligan & Cooper, 1985; Sarle, 1983). Positive values of CCC greater than 2.0 "mean that the obtained R^2 is greater than would be expected if sampling from a uniform distribution, and therefore, indicate the possible presence of clusters" (p.4, Sarle, 1983). The CCC statistic tests the alternative hypothesis that: "the data have been sampled from a mixture of spherical multivariate normal distributions with equal variances and equal sampling

[2]The k-means clustering procedure can not be used with a distance matrix, so we did not perform this analysis for Project A.

probabilities" (p.4, Sarle, 1983). Milligan and Cooper (1985) compared CCC to 29 other stopping rules and found that it was quite successful in recovering the proper number of clusters from synthetic data sets, where the true number of clusters was known--performing sixth best.

A second index, the standardized version of Hubert's (Hubert & Arabie, 1985) Gamma, was also used to determine the number of clusters. In standardized form Gamma is the sample correlation between the entries of two matrices, in this case a cluster solution and a distance matrix (Jain & Dubes, 1988). The numerator is the difference between consistent cluster memberships and inconsistent memberships for all pairs of objects. A consistent pair of objects occurs when objects that are assigned to the same cluster have smaller distances than objects assigned to different clusters. Inconsistency occurs when objects in the same cluster have larger distances than objects in different clusters. The denominator is the total number of object pairs, or n(n-1/2), where n is the number of objects. Gamma ranges in value from -1 to +1, and is corrected for chance matches in the two matrices. Gamma is 1.0 when a cluster solution is perfectly consistent with the underlying data matrix, and 0.0 when pairs match by chance.

We also examined the usefulness of the point-biserial correlation for choosing cluster solutions, because it is a familiar measure to industrial psychologists. The point-biserial correlation is a raw measure of agreement between a dichotomous and continuous variable (Jain & Dupes, 1988; Lord & Novick, 1968). The range of the coefficient varies with the ratio of 1's to 0's on the dichotomous variable (i.e., the cluster solution matrix). Milligan and Cooper (1985) found it to be only slightly poorer than CCC in selecting the number of clusters. However, we decided that it was not suitable for comparing values across different numbers of clusters, because the ratio of 1's to 0's would probably vary quite a bit. We tested this and found it to be true.

Procedure

The procedure for clustering the data was to examine a large number of solutions for each clustering algorithm. The cluster structures differed in the number and composition of the clusters. The objective of the internal validation was to select the solution with the fewest clusters that provided an adequate fit with the underlying distance matrix. Cluster solutions containing between 2 and 50 clusters were examined for the DOT and Project A jobs. Solutions with 2 to 20 clusters were obtained for the 21 synthetic validity jobs, using both the CTI and OJI task ratings. SAS provides CCC values when the number of clusters is 20 percent or less of the number of jobs. Therefore, CCC values were computed for solutions with 2, 3 and 4 clusters in the synthetic validity data, and for up to 50 clusters in the other two data bases. Gamma was computed for 2 to 50 clusters for the two large data sets and 2 to 20 clusters for the synthetic validity data. The CCC and Gamma values were then plotted against the number of clusters for each of the three algorithms: Ward, average linkage, and K-means.

The CCC plots were evaluated by looking for large jumps in values, or for peaks in the distribution, as the number of clusters increased (Sarle, 1983). Gamma was assessed by looking for peaks. When the indexes selected different cluster solutions, CCC was used to weed out potentially random solutions. The final decision was made by selecting the single best solution

within the nonrandom set using the Gamma plot. If several solutions had about the same Gamma value, then the solution with the smallest number of clusters was selected. We used CCC in combination with Gamma in our study because CCC provided a means for evaluating statistical significance, while no test was available for Gamma, the better measure in Monte Carlo studies.

Previous studies based on the four data sets enabled us to make predictions about the expected number of clusters. Harris et al. (1991) obtained 13 job families, when they clustered all military jobs using the DOT data. We expected about the same number, or slightly fewer, since our study examined only Army jobs. Both Rosse et al. (1983) and Hoffman (1987) obtained 23 clusters in the Project A data. The synthetic validity data base of 21 jobs represents 16 of the 23 Project A clusters (Hoffman, 187). As mentioned above, Peterson et al. (1991) used the Ward HCA procedure, with correlations as proximity measures, to cluster the OJI and CTI data. They obtained 4 clusters for CTI and 3 clusters for OJI. No hypotheses were proposed for differences in the clustering algorithms, except that we expected the Ward and K-means solutions to be fairly similar because they both minimize within-cluster variance.

Cluster Structure Consistency[3]

Two types of consistency analyses were conducted: 1) comparisons of the results of alternative clustering procedures, and 2) cross-validation. The objective was to assess whether the composition of the groups would be stable across algorithms and cross-validation samples. If we found consistency, despite not being able to conduct statistical tests, then we would have some confidence that the cluster structure was robust.

We used the simple matching coefficient attributed to Rand (1971) to compare pairs of cluster structures in both sets of consistency analyses. The Rand is the sum of consistent comparisons across two cluster structures, represented as binary variables for the same set of objects. The value is 0.00 when two jobs are in the same cluster and 1.00 when they are in different clusters. Consistent matches between two cluster solutions are found when: 1) two jobs are in the same clusters in the two solutions, <u>or</u> 2) the jobs are in different clusters in the two solutions. The numerator of the Rand is the sum of the two types of consistent comparisons and the denominator is the total number of pairs of jobs. The range of the Rand is 0.0 to 1.0.

Hubert and Arabie (1985) demonstrated that the Rand statistic is a special case of Gamma. Previous research indicated that the raw Rand produced inflated values with little variance (Jain & Dubes, 1988; Milligan & Schilling, 1985). For example, Fraboni and Cooper (1989) used the raw Rand statistic to compare six hierarchical clustering algorithms for grouping the subtests of the Wechsler Adult Intelligence Scale-Revised. They obtained only three values for 15 comparisons: .49, .67 and 1.00. Hubert and Arabie (1985) developed a formula for correcting the raw Rand for chance matches, which should result in more realistic estimates and increased variance. An earlier correction formula by Milligan and Schilling (1985) was found to be incorrect. We calculated both the corrected (Rand_c) and raw Rand.

[3]The synthetic validity data bases were dropped from these and the remaining analyses because of the small sample size.

Comparison of Clustering Procedures

The first set of consistency analyses involved comparing clustering algorithms. The $Rand_c$ index was computed for comparisons between the Ward and average linkage, Ward and K-means, and average linkage and K-means, procedures.

Cross-validation

A double cross-validation analysis was conducted with the DOT jobs.[4] We modified a procedure proposed by McIntyre and Blashfield (1980). Their approach was to randomly divide a data set in half and to cluster Sample One. Then they used the cluster centroids from Sample One to form clusters in Sample Two. Sample Two was then clustered directly, and the two clusterings of Sample Two were compared using a measure of agreement. We followed this procedure for both samples, and then took the average Rand as our estimate of stability, thus performing a double cross-validation analysis.

We hypothesized that the two sets of stability analyses would show consistent cluster structure across clustering algorithms and cross-validation samples.

External Comparisons

We conducted two types of external comparisons. Although we could not compare our empirical solutions to the true cluster structures of the data bases, we did compare them to the Army Aptitude Areas (AA). As discussed in Chapter One, the AA were formed by grouping jobs with similar patterns of validity coefficients on the ACB, and through SME evaluations of the similarity of performance requirements. Our research question in the first set of external analyses was whether any of the job family structures, based on different types of job descriptors and three different empirical clustering algorithms, would be similar to the operational AA job families, derived by a different process. We anticipated that the AA would show little overlap with the cluster solutions obtained in this study.

The final step in the evaluation method was to compare the cluster structures obtained in the DOT and Project A data sets for the 227 overlapping jobs. We asked the same general question: Are the cluster solutions obtained from different types of job analysis data for the same set of jobs consistent? We expected little overlap in the job families.

[4] The cross-validation required average scores on the descriptors for each job, and consequently, could not be conducted for the Project A data base, which consisted of only a distance matrix.

CHAPTER III

RESULTS

Appendix A contains the MOS in the four data bases we studied, and the Aptitude Areas (AA) and Career Management Fields (CMF) to which they belong. Appendices B through L present the results of the cluster analyses for the three clustering procedures in the four data bases.

Internal Validity

The objective of the internal validity analysis was to select the single best solution from among a large number of solutions produced by each clustering algorithm. Appendices B through D present the results of the Ward, average linkage and K-means cluster analyses of the DOT jobs, respectively. Figures 1 and 2 in each appendix contain plots of the CCC and Gamma values by number of clusters. Table 1 in the appendices contains the assignments of jobs to clusters, the job title and the distance of each job from the cluster centroid. The average within-cluster distances and mean cluster factor scores are also presented.

The results of the Ward and average linkage clusterings of the Project A jobs are presented in Appendices E and F. Figure 1 contains the plot of Gamma by number of clusters. CCC could not be obtained for the Project A data base, which consisted of a distance matrix, because this index is calculated from mean scores on a set of variables for each job. Although the form of the Project A data limited our analyses, we decided to retain it in the study, because it contained a large number of MOS, many of which overlapped with the DOT jobs. This would allow us to make comparisons between the two data bases, which contain different types of job descriptors. Table 1 in Appendices E and F presents the cluster assignments.[5] Appendices G through I contain the three cluster solutions for the synthetic validity OJI data, and Appendices J through L present the results for the CTI data.

Table 4 below shows the number of clusters that were selected by examining the CCC and Gamma plots for each clustering method and data base. The number of clusters varied across clustering algorithms and data bases. In general, the largest number of clusters was obtained with the average linkage algorithm, and in the Project A data base. However, the average linkage procedure produced fewer clusters than the Ward and K-means procedures in the synthetic validity CTI sample.

The differences in the number of clusters found for the DOT and Project A data bases were in line with expectations. We anticipated a relatively small number of clusters in the DOT data, given the global nature of the factor scores, and the Harris et al. (1991) finding of 13 clusters for 900 military jobs. We found about twice as many clusters in the Project A data base, and close to the 23 clusters obtained by Rosse et al. (1983) and Hoffman (1987).

---

[5]Note that distance and mean factor scores could not be computed for the Project A jobs because we did not have job descriptor scores on individual observations.

Table 4. Number of Clusters Selected by Three Cluster Analysis Methods for DOT, Project A and Synthetic Validity Data Bases

Data Base	Cluster Analysis Method: Ward	Average Linkage	K-Means
DOT	7	13	6
Project A	17	21	-
Synthetic Validity-OJI	8	10	10
Synthetic Validity-CTI	13	9	17

The cluster analyses of the synthetic validity data produced different numbers of clusters for OJI and CTI ratings. Examination of the Gamma plots identified peaks for structures with between 8 and 10 clusters for OJI ratings, and 9 to 17 for CTI.[6] These results are quite different from the cluster solutions obtained in earlier studies. Recall that the synthetic validity jobs are a sample from the Project A data base. The 21 MOS represent 16 of the 23 Hoffman (1987) clusters. Recall also that Peterson et al. (1991) found 3 clusters for OJI and 4 clusters for CTI.

The remaining analyses explore the similarities and differences in the cluster structures of the DOT and Project A data bases. The synthetic validity data were dropped from further analysis because of the small sample size, which probably contributed to the inconsistencies in findings across studies.

Consistency Analyses

Comparison of Clustering Procedures

Tables 5 and 6 below present the corrected and raw Rand values comparing pairs of algorithms used to cluster the DOT and Project A data bases, respectively. Examination of the corrected Rand values ($Rand_c$) shows that the cluster structures produced by the three procedures are moderately consistent for both the DOT and Project A data, despite the variation in the number of clusters.

For the DOT data, the greatest consistency was found between the Ward and K-means clusters ($Rand_c$ = .6860), and the least consistency was found between the average linkage and K-means structures ($Rand_c$ = .5207). As mentioned above, the Ward and K-means procedures are similar in that they both use the same optimizing criterion--minimum within-cluster variance. They differ in that the Ward procedure produces a hierarchical nesting of job clusters, while the K-means algorithm does not. Further, the K-means

[6]The CCC plots contained in Appendices G through L show that solutions with 2 to 4 clusters were not significantly different from chance in our analyses. Unfortunately, CCC values are not computed by the SAS cluster analysis program, when the number of clusters is greater than 20 percent of the objects. Therefore, we could not evaluate other solutions using this index.

Table 5. Rand Values Comparing DOT Cluster Structures Across Cluster Analysis Methods[a]

Cluster Analysis Method and Number of Clusters	Ward (7)	Average Link (13)	K-Means (6)
Ward (7)	1.00	.6200 (.9011)	.6860 (.9137)
Average Link (13)		1.00	.5207 (.8620)
K-Means (6)			1.00

[a]First value is $Rand_c$, value in parentheses is raw Rand.

Table 6. Rand Values Comparing Project A Cluster Structures Across Cluster Analysis Methods[a]

Cluster Analysis Method and Number of Clusters	Ward (17)
Average Link (21)	.7087 (.9681)

[a]First value is $Rand_c$, value in parentheses is raw Rand.

procedure was designed to improve upon the Ward method by reassigning jobs after initial clustering to improve homogeneity. The results show that reassignment did not make much difference in the cluster structure.

Table 7 presents a more detailed picture of the similarities and differences between the cluster structures produced by the Ward and K-means algorithms, and is the contingency table that forms the basis for computing the $Rand_c$ index. The columns contain the number of jobs in the K-means solution. The rows contain the number of jobs in the Ward clusters. The elements are the number of jobs in a particular cluster of each solution that overlap. Completely consistent solutions will have one non-zero value in each row (or column), with the remaining elements being zero. High to moderate agreement is found when one or two of the clusters of one solution are split into a small number of clusters in the other solution, leaving a relatively large number of zeros in the rows (or columns). Low agreement is seen when the cluster members of one solution are spread throughout a large number of clusters in the other solution, leaving few zero elements.

Examination of Table 7 shows that most of the assignments were consistent in the Ward and K-means structures of the DOT jobs. The major differences are seen in the assignments in K-means clusters 4 and 5 and Ward clusters 1, 3, and 6.

If a set of job families were being developed for research or operational purposes, this Rand contingency table would be useful for understanding the differences in the two cluster structures and for

Table 7. Rand Contingency Table Comparing DOT Ward and K-Means Solutions

Ward Clusters	K-Means Clusters CL1	CL2	CL3	CL4	CL5	CL6	Row Sums
CL1	2	36	0	8	0	0	46
CL2	24	0	0	4	0	0	28
CL3	0	3	4	0	19	0	26
CL4	0	0	67	0	0	0	67
CL5	0	0	0	43	0	0	43
CL6	1	0	0	15	13	0	29
CL7	0	1	0	0	0	23	24
Column Sums	27	40	71	70	32	23	263

determining the final cluster memberships. For example, the fit of the discrepant jobs in the inconsistent clusters of the two structures could be evaluated through both qualitative and quantitative analysis. The qualitative analysis would rest upon comparing the congruence of the titles and performance requirements of the discrepant jobs with the other jobs in the clusters of each solution. Quantitative analysis would include evaluating the distance of each job from its cluster centroid in the two structures, and comparing each job's factors scores to the cluster average scores in the alternative solutions. The final assignments would place the jobs in the most similar cluster using researcher judgement or validation against an external criterion, e.g., classification efficiency.

Tables 8 and 9 present the Rand_c contingency tables comparing the Ward and average linkage, and average linkage and K-means procedures, respectively, in the DOT data base. These comparisons also show fairly high overlap. The average linkage algorithm was not quite as consistent with the Ward method as the K-means procedure. Although they are both hierarchical algorithms, average linkage and Ward use different optimizing criteria for cluster assignment. The least similar solutions, average linkage and K-means, differ in whether the algorithm is hierarchical or non-hierarchical and in the optimizing criterion.

Further examination of Tables 7, 8 and 9 shows that the Ward and K-means algorithms tended to produce clusters of about the same size, (between 23 and 71 jobs); while the average linkage method produced a large number of big clusters and a few small clusters with between 1 and 7 members. This pattern was also found for Project A jobs. Table 10 contains the Rand_c contingency table for the Ward and average linkage solutions in the Project A data, which also shows high overlap.

Table 8. Rand Contingency Table Comparing DOT Ward and Average Linkage Solutions

Ward Clusters	Average Linkage Clusters													
	CL 1	CL 2	CL 3	CL 4	CL 5	CL 6	CL 7	CL 8	CL 9	CL 10	CL 11	CL 12	CL 13	Row Sums
CL1	33	0	1	0	0	12	0	0	0	0	0	0	0	46
CL2	0	20	0	0	0	0	0	7	0	0	0	0	1	28
CL3	0	0	18	0	0	0	0	0	4	0	4	0	0	26
CL4	0	0	0	67	0	0	0	0	0	0	0	0	0	67
CL5	0	0	0	21	22	0	0	0	0	0	0	0	0	43
CL6	0	1	0	0	12	0	16	0	0	0	0	0	0	29
CL7	4	0	0	0	0	0	0	0	0	18	0	2	0	24
Column Sums	37	21	19	88	34	12	16	7	4	18	4	2	1	263

Table 9. Rand Contingency Table Comparing DOT Average Linkage and K-Means Solutions

Average Linkage Clusters	K-Means Clusters						
	CL1	CL2	CL3	CL4	CL5	CL6	Row Sums
CL1	0	26	0	8	0	3	37
CL2	21	0	0	0	0	0	21
CL3	0	4	4	0	11	0	19
CL4	0	0	67	21	0	0	88
CL5	0	0	0	34	0	0	34
CL6	2	10	0	0	0	0	12
CL7	0	0	0	3	13	0	16
CL8	3	0	0	4	0	0	7
CL9	0	0	0	0	4	0	4
CL10	0	0	0	0	0	18	18
CL11	0	0	0	0	4	0	4
CL12	0	0	0	0	0	2	2
CL13	1	0	0	0	0	0	1
Column Sums	27	40	71	70	32	23	263

Table 10. Rand Contingency Table Comparing Project A Ward and Average Linkage Solutions

Ward Clusters	Average Linkage Clusters										
	CL1	CL2	CL3	CL4	CL5	CL6	CL7	CL8	CL9	CL10	CL11
CL1	17	0	0	0	0	0	0	0	0	0	0
CL2	0	7	0	0	0	0	0	0	0	0	0
CL3	0	0	13	0	0	0	0	0	0	0	0
CL4	0	0	0	23	0	0	0	0	0	0	0
CL5	0	0	0	0	11	0	0	0	0	0	0
CL6	0	0	0	0	0	12	0	0	0	0	0
CL7	0	1	0	0	1	2	3	0	0	0	0
CL8	0	0	0	0	0	0	0	20	0	0	0
CL9	0	0	0	0	0	0	0	0	9	0	0
CL10	0	0	0	0	0	0	0	0	0	14	0
CL11	0	0	0	0	0	0	0	0	0	0	23
CL12	0	0	0	0	0	0	0	0	0	0	0
CL13	0	0	0	0	0	0	0	0	0	0	0
CL14	0	0	0	0	0	0	0	0	0	0	1
CL15	0	0	0	0	0	0	0	0	0	0	0
CL16	0	0	2	0	0	0	0	0	0	0	4
CL17	2	0	0	0	0	0	0	0	0	0	4
Column Sums	19	8	15	23	12	14	3	20	9	14	32

Table 10 Con't. Rand Contingency Table Comparing Project A Ward and Average Linkage Solutions

Ward Clusters	Average Linkage Clusters										
	CL12	CL13	CL14	CL15	CL16	CL17	CL18	CL19	CL20	CL21	Row Sums
CL1	0	0	0	0	0	0	0	0	0	0	17
CL2	0	0	0	0	0	0	0	0	0	0	7
CL3	0	0	0	0	0	0	0	0	0	0	13
CL4	0	0	0	0	0	0	0	0	0	0	23
CL5	0	0	0	0	0	0	0	0	0	0	11
CL6	0	0	0	0	0	0	0	0	0	0	12
CL7	0	0	0	3	3	6	4	3	0	2	28
CL8	0	0	0	0	0	0	0	0	0	0	20
CL9	0	0	0	0	0	0	0	0	0	0	9
CL10	0	0	0	0	0	0	0	0	0	0	14
CL11	0	0	0	0	0	0	0	0	0	0	23
CL12	16	0	0	0	0	0	0	0	0	0	16
CL13	0	10	0	0	0	0	0	0	0	0	10
CL14	0	0	31	0	0	0	0	0	0	0	32
CL15	0	0	0	6	0	0	0	0	0	0	6
CL16	0	0	7	0	0	0	0	0	0	0	13
CL17	5	0	0	0	0	0	0	0	3	0	14
Column Sums	21	10	38	9	3	6	4	3	3	2	268

Double Cross-Validation

Figure 1 and Table 11 present the results of the double cross validation analysis of the DOT data. Figure 1 shows that the number of clusters obtained in each sample was different, and differed from the number obtained when the total population of jobs was clustered. The Ward and K-means procedures produced between 4 and 7 clusters, depending on the sample, and the average linkage procedure produced between 11 and 13 clusters. The intersections of the lines for the two samples and the population suggest there is an interaction between the sample and the clustering procedure.

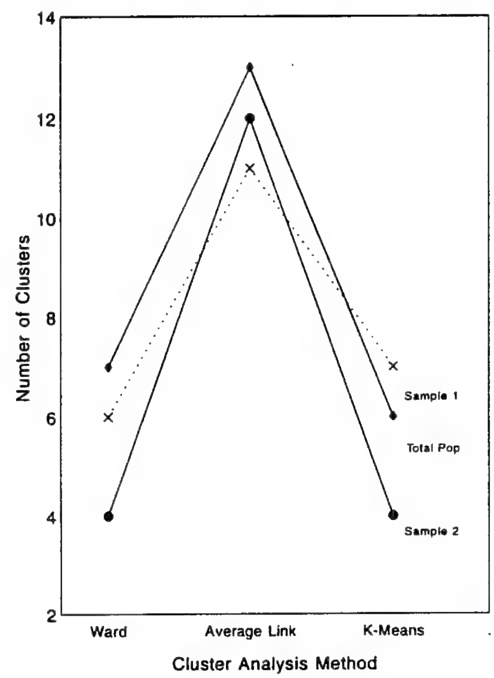


Figure 1. Number of Clusters in DOT Population and Cross Samples One and Two by Cluster Analysis Method

Table 11 contains the average $Rand_c$ values comparing the direct and indirect cluster solutions of cross-validation Samples One and Two. The results show low to moderate overlap in the structures. The lowest level of agreement in Samples One and Two was found for the K-means solutions. Examination of the $Rand_c$ contingency table (not shown), from which the $Rand_c$ value of .2447 was calculated, showed that there was a great deal of spread in the assignments across the K-means solutions. On the other hand, the Ward ($Rand_c$ = .4272) and average linkage solutions ($Rand_c$ = .5702) tended to have a large number of consistent classifications. When there were discrepancies, the jobs in a cluster in one sample tended to fall into two large groups, or into one large and one small group, in the other sample.

Table 11. Rand Values Comparing Sample One and Sample Two Cluster Solutions in DOT Double Cross-Validation Analysis[a]

Cluster Analysis Method	Rand Values
Ward	.4272 (.7761)
Average Linkage	.5702 (.8940)
K-Means	.2447 (.7173)

[a]First value is $Rand_c$, value in parentheses is raw Rand.

External Comparisons

Two types of external comparisons were conducted. In the first, the DOT (total sample), and Project A cluster solutions were compared to the Army AA. The Rand values are shown in Tables 12 and 13, respectively. The overlap is quite poor, with $Rand_c$ values of .2671 for K-means, .3225 for average linkage, and .3324 for Ward in the DOT data, and .2475 for Ward and .2839 for average linkage in the Project A data. Table 14 provides a more detailed picture of the extent of differences between one of the cluster structures (Ward solution for DOT jobs), and the AA. The contingency table shows that the jobs in the nine AA are quite spread out across the empirical solution--there are relatively few zeros in any of the columns. This pattern is repeated in all the other contingency tables (not shown).

Table 12. Rand Values Comparing DOT Empirical Cluster Solutions with Nine Army Aptitude Area Job Families[a]

Cluster Analysis Method and Number of Clusters	Rand Values
Ward (7)	.3324 (.8110)
Average Linkage (13)	.3225 (.8048)
K-Means (6)	.2671 (.7755)

[a]First value is $Rand_c$, value in parentheses is raw Rand.

Table 13. Rand Values Comparing Project A Empirical Cluster Solutions with Nine Army Aptitude Area Job Families[a]

Cluster Analysis Method and Number of Clusters	Rand Values
Ward (17)	.2475 (.8380)
Average Linkage (21)	.2839 (.8442)

[a]First value is $Rand_c$, value in parentheses is raw Rand.

The second set of external comparisons measured the congruence of the DOT and Project A structures using the 227 overlapping jobs (see Table 15). The $Rand_c$ values of .1671 for the Ward clusters and .2143 for the average linkage clusters are also quite low. These results suggest that the type of job descriptor has a substantial impact on the composition of the clusters. However, differences in the samples of jobs also undoubtedly contribute to the low agreement.

Table 14. Rand Contingency Table Comparing DOT Ward and Army Aptitude Area Job Families

Ward Clusters	Army Aptitude Areas									
	CL	CO	EL	FA	GM	MM	OF	SC	ST	Row Sums
CL1	14	1	4	0	5	2	2	3	15	46
CL2	0	8	4	1	6	2	4	1	2	28
CL3	2	1	0	2	1	0	0	1	19	26
CL4	0	0	57	0	4	0	0	1	5	67
CL5	0	0	11	0	4	26	0	0	2	43
CL6	1	2	3	0	20	3	0	0	0	29
CL7	3	0	1	0	0	0	3	0	17	24
Column Sums	20	12	80	3	40	33	9	6	60	263

CL=Clerical
CO=Combat
EL=Electronics Repair
FA=Field Artillery
GM=General Maintenance
MM=Mechanical Maintenance
OF=Operators and Food
SC=Surveillance and Communications
ST=Skilled Technical

Table 15. Rand Values Comparing DOT and Project A Cluster Structures[a]

Cluster Analysis Method	Rand Values
Ward	.1671 (.8348)
Average Linkage	.2143 (.8396)

[a]First value is $Rand_c$, value in parentheses is raw Rand.

CHAPTER IV

DISCUSSION AND CONCLUSIONS

The test of the proposed evaluation method suggests that it is a good validation and diagnostic tool for applied clustering problems. It can be used with any clustering procedure, ranging from rationally-derived clusters to factor analysis and hierarchical cluster analysis. Although it assumes that clusters are mutually exclusive, it can be used to identify overlapping job families.

Internal Validity

One of the most difficult tasks in cluster analysis is to select the number of clusters that provides the best representation of the underlying relationships among a set of objects, in our case jobs. This is especially difficult in applied research when the true cluster structure is unknown. Milligan and Cooper (1985) examined a large number of internal validity indexes with synthetic data where the true number of clusters was known. Two of the best were Gamma and CCC. Each measure uses a different approach for measuring the quality of a cluster structure. They found that when the indexes were inaccurate, Gamma tended to select solutions with too few clusters, while CCC selected solutions with too many clusters. Since the true cluster structures of our data are unknown, we could not compare the accuracy of the two indexes.

Gamma demonstrated valuable properties as an internal validity index in this study. The values varied by the number of clusters, reaching a discernable peak, and either descended as the number of clusters increased, or remained at about the same level. In contrast, the CCC plots were fairly difficult to interpret. The CCC values generally increased as the number of clusters increased, showing no peaks. Therefore, the best cluster solution according to CCC was identified by looking for relatively large jumps in values greater than 2.0. The major advantage of the CCC index is its statistical interpretation. Research is needed to develop a statistical significance test for Gamma. Two potential approaches are discussed below under Future Research.

Comparison of the Gamma and CCC plots for each cluster solution in Appendices B through D and G through L found that they were consistent for the Ward and K-means solutions, but that the indexes were somewhat conflicting for the average linkage solutions. It may be that the method of combining jobs into clusters used in the average linkage procedure (i.e., comparing the average distance between all pairs of jobs in each cluster), cannot be accurately evaluated by one or both of the indexes. Further research is needed to understand the relationship between the clustering criterion and the index of internal validity.

The results of the cluster analyses showed that the number of groups varied across data base and clustering algorithm. The numbers obtained for the DOT and Project A data bases were fairly consistent with the results of previous studies of the same data. The relatively large number of clusters we obtained for the synthetic validity OJI and CTI data differed from the results of Peterson et al. (1991), who found 3 and 4 clusters, respectively.

Cluster Structure Consistency

Comparison of Clustering Methods

The comparisons of clustering algorithms produced four major findings. First, although the number and composition of the clusters differed somewhat with the clustering procedure, the jobs in both the DOT and Project A data bases had fairly stable cluster structures. If there were no clusters in the data, then the results of each procedure would be random and the expected $Rand_c$ values would be zero.

Second, the rank ordering of the $Rand_c$ values for the DOT cluster solutions shown in Table 5 suggests that the optimizing criterion used in a clustering procedure had a greater impact on the similarity of two cluster structures, than whether the algorithm was hierarchical or non-hierarchical.

Third, Harvey (1986, 1991) hypothesized that the structure of job families is overlapping, rather than mutually exclusive. Given this hypothesis, the differences we found across clustering algorithms may be due, at least in part, to jobs that fit in more than one family. Procedures such as factor analysis and multidimensional scaling should be considered in future analyses of these and other Army job analysis data, because these clustering methods show the relationships of jobs to all groupings and do not force them into mutually exclusive clusters.

Fourth, as expected, examination of the raw Rand values in all relevant analyses (comparisons of clustering procedures, cross-validation analysis and external comparisons), found a pattern of inflated values with low variance. The range for the corrected index was .1671 to .7087 across all analyses, while the range of the raw Rand was .7173 to .9681.

The major limitation inherent in comparing cluster solutions of the same objects, when working with real data, is that we cannot know the true cluster structure. Consequently, we cannot determine which clustering procedure provides the best fit. In addition, we cannot determine the extent to which the differences in the results are attributable to the methods or to the structure in the data, e.g., overlapping clusters. Further analysis of the composition of the clusters, and of outliers, would provide more information about the possible sources of variation in the cluster results.

In summary, the findings support the overall conclusion from the job family research reviewed in Chapter One that the choice of a clustering method is a complex decision. Garwood et al. (1991) found that the accuracy of different procedures depends on the selection of the job sample, the structure of the organization, and the purpose for which jobs are being clustered. They suggested, and the results of the present study support, a strategy of using several different clustering methods. The discrepancies across procedures should be resolved by expert judgement or external validation.

Double Cross-Validation

The purpose of the cross-validation analysis was to measure the extent to which the cluster results produced by each procedure could be replicated in different samples from the same population. If the cluster structure is

stable across different samples, then we can be confident we obtained the true cluster solution. The disadvantage of cross-validation is that it requires the original sample (or population in our case) to be randomly divided in half, thereby, increasing the effect of sampling error.

The absence of stability across the DOT samples was disappointing, and suggests that the sample is highly important in clustering job families. We can only speculate about the causes of variation. The fairly good stability seen in the first set of consistency analyses, i.e., across clustering procedures, indicates some stability in cluster structure. The discrepancies found in both sets of consistency analyses may be due to overlapping job families, or to the presence of outliers--jobs that do not fit well in any family and should be considered single-job families. Both situations would cause these jobs to be randomly clustered by different procedures and in different samples, thus producing low $Rand_c$ values. Another potential cause of the low consistency in the cross samples could be that the clustering algorithms we examined are highly sensitive to the sample. This may be especially a problem for the K-means and Ward procedures, both minimum variance techniques, which have the lowest cross-validation $Rand_c$ values.

Given these results, we recommend that every attempt be made to obtain data on the total population of jobs, or that the sample be selected with great care when forming job families. Both Rosse et al. (1983) and Hoffman (1987) clustered the Project A data base. They obtained high overlap in the solutions, although Rosse et al. had a sample of 111 MOS, and Hoffman clustered the total population. The consistency of their results was probably, at least, partly due to the use of a factor analytic procedure that identifies jobs that fit into more than one job family, and relies on researcher judgement to make the job family assignments. The three procedures we studied were purely empirical approaches. When job families are developed for research or operational purposes, the results of clustering algorithms should be evaluated using expert judgement. Further, in an operational study, the cluster structure must be evaluated in terms of its effectiveness for satisfying an external criterion.

Another approach to evaluating the stability of job clusters through cross-validation analysis would be to obtain multiple ratings on the job descriptors within each job. First, the sample would be randomly divided in half within the jobs, forming two samples containing all jobs, but with one-half the number of respondents in each sample. Then each sample would be clustered. The $Rand_c$ index computed between the two samples would be the estimate of consistency. If the results were fairly stable, then the samples could be combined, and a final set of job families developed using all of the data. This approach was not possible in the present study because we had only mean scores on the variables in each data base. However, it would fit quite nicely with the typical job analysis procedure that obtains ratings on job descriptors from a sample (or the population) of job incumbents.

External Comparisons

Comparison of Empirical Clusters with Army AA

Little congruence was found between the empirical cluster structures and the AA job families. The results were equally poor for the three clustering algorithms and the two data bases. Without conducting an external validation

study, in which the empirical and existing AA clusters are evaluated against an external criterion, we cannot determine which set of clusters, if any, is effective for a particular purpose. The AA are used to reduce the complexity of the Army's personnel classification system, and any clusters designed to replace them should be evaluated in terms of their classification efficiency.

Two studies (Johnson et al.,1992; Statman, 1993) indicate that the AA do not provide much classification efficiency. A small sample of MOS--18 of the 21 jobs in the synthetic validity study, were clustered using different approaches. Johnson et al. developed a hierarchical cluster analysis procedure that minimized the reduction in the differential validity of the set of prediction equations for the jobs as they were grouped into clusters. Statman used a factor analytic approach in which the prediction equations for the jobs were factored, and jobs were assigned to clusters according to their loadings on orthogonal factors. The results in both studies showed that increasing the number of job families, and configuring them to maximize differential validity, provided substantial increases in classification efficiency.

The DOT and Project A clusters obtained in the present study were derived by different procedures, and based on different types of job information compared to the classification research. The low $Rand_c$ values provide limited support for the earlier classification studies in that the empirical clusters bear little relationship to the AA. Taken together, the present results, and those of Johnson et al. (1992) and Statman (1993), suggest that the number and composition of the AA should be reevaluated.

Comparison of DOT and Project A Clusters

The external comparisons between the DOT and Project A cluster structures found little congruence in both the Ward and average linkage cluster structures. The very low values suggest that the type of data (DOT factor scores vs. Project A SME groupings), has a strong effect on cluster structure. Given the need to develop job families that are efficient in accomplishing one or more organizational objectives, the type of job descriptor should reflect the important aspects of the personnel process for which the clusters are designed (Sackett, 1988).

More research is needed on the relationship of different job descriptors to alternative personnel processes. Selection has received the most attention. Two studies described in Chapter One (Cornelius et al., 1984; Sackett et al., 1981), found that global job descriptors provide about the same cluster structure as more specific task descriptors. However, Stutzman (1983) demonstrated that task differences can be important in differentiating job families in some situations. Cornelius et al. (1979) also observed that job families developed for jobs requiring training after selection should be based on different descriptors than jobs that require immediate performance. Overall, there are no well researched guidelines at present on the choice of the variables or clustering procedures for constructing job families for different purposes.

Limitations of the Research

There were several limitations of the present study. The first was the inability to conduct hypothesis testing. At present cluster analysis is an exploratory data analysis technique. A cluster structure is usually selected by subjective evaluation of the results on the part of the researcher by examining a tree diagram, assessing within and between cluster distances, or conducting a content analysis.

The proposed cluster evaluation method used the Gamma and CCC indexes to select cluster structures. The Rand index was used as a measure of the agreement between two cluster solutions. Unfortunately, the sampling distributions of Gamma, the $Rand_c$ index, and other measures of internal and external cluster validity, are unknown. Therefore, it was not possible to conduct statistical hypothesis testing, except for the internal validity analysis when the CCC statistic was available.

This limitation was partially overcome by the two types of consistency analyses we conducted. The comparisons across clustering procedures, and the cross-validation analysis, provided a non-statistical method (i.e., replication analysis), for assessing whether the true cluster structure was recovered. However, replication analysis is limited in that the sources of error cannot be identified when there is little or no agreement between a pair of cluster structures (Milligan & Cooper, 1987). An external validation is needed in applied research settings to evaluate cluster structure against a relevant criterion.

Another limitation of the study was that no outlier analyses were conducted to identify jobs that did not fit well into the clusters of any solution. Previous research (Jain & Dubes, 1988) indicates that excluding them from the analysis, or sometimes considering them as separate clusters, often improves the fit of the overall solution. Future application of the evaluation methodology should include outlier analysis.

Lastly, no attempt was made to subjectively evaluate the differences in the cluster solutions obtained by the Ward, average linkage and K-means procedures. The cluster comparison analyses found that, although the three algorithms produced overlapping structures, there were discrepancies. If the evaluation method were used to develop a set of operational clusters, an important part of the development process would be to resolve the differences in the procedures through either expert judgement or external validation.

Future Research

One of the most important unresolved problems in cluster analysis is hypothesis testing (Sokal, 1988). Further research is needed to develop statistical procedures for evaluating the significance of the quantitative indexes examined in this study. Two approaches have been suggested (Jain & Dubes, 1988; Milligan & Cooper, 1987). The first is to use Monte Carlo procedures to generate a sampling distribution of multivariate random data having the same number of variables, means, variances and covariances as the empirical data. This is the method used with the CCC statistic we included in the internal validity analysis. A second approach is to use a permutation model, in which the observed value of an internal or external index is compared to a sampling distribution of 100 to 200 random cluster solutions of

the data. Neither procedure has received much attention in the applied literature. Consequently, important questions (e.g., the appropriate random distribution for different data sets generated by a Monte Carlo procedure, and the relevant model of randomness for a permutation procedure), must be answered before statistical procedures can be developed.

A second area of future research is the evaluation of the effectiveness of the Army AA and the CMF. The results of the external comparisons of AA with the empirical cluster solutions (combined with the previous research of Johnson et al., 1992, and Statman, 1993), indicate that the number and composition of the AA is probably not optimal for classification. A cluster analysis study that includes the following components is needed: 1) investigation of several different clustering strategies, 2) analysis of alternative types of job descriptors, and 3) an external validation.

Given the structural changes currently taking place in MOS, and in the Army organization-wide, the CMF also should be reevaluated, and reconstructed, if necessary, to ensure their utility for force management, and career development and training. Examination of the range of CMF within the empirical cluster solutions in Appendices B through L found that the clusters were highly heterogeneous. Further research is needed to evaluate whether the clusters found in this study would be more effective than the existing CMF.

Beyond job family research, the proposed cluster evaluation method would make an important contribution to any cluster analysis procedure by providing a quantitative basis for evaluating and validating cluster structures. A third area of future research is the use of the proposed method for clustering tasks, or other job descriptors, into jobs. Working at the job level, the method could be used to validate the basis for restructuring MOS during the current downsizing.

An evaluation of existing tasks clusters, or development of new clusters, should examine the consistency of several different approaches. The finding that the average linkage algorithm produced more clusters than either the Ward hierarchical or the K-means procedures (both minimum variance techniques), indicates that the procedures are defining clusters in different ways. The results should be evaluated against an external criterion wherever possible to determine the best set of clusters. External validation may be difficult when designing new MOS and restructuring old ones. However, some attempt should be made to develop an external validation strategy within the practical constraints of project resources and time. Ratings of alternative cluster structures by SMEs, or work sample tests of the ease or difficulty of accomplishing the tasks in the clusters within a given time, are two possible approaches.

A double cross-validation analysis also should be included in an MOS restructuring project to evaluate the stability of the cluster structures. If the clusters are stable in different samples, then they probably represent natural groupings of tasks. If the results are unstable, then the discrepancies between the samples should be evaluated. If a set of tasks appear to fall randomly into clusters in different samples (or with different procedures), then this could indicate the absence of a true cluster structure or the presence of overlapping groups. The assignments of the discrepant tasks could then be handled through a content analysis of the solutions and by considering practical issues.

The last component of an MOS restructuring study should include external comparisons of the new task clusters with existing ones using the $Rand_c$ index, as we did in this study. This analysis would be a good diagnostic tool for assessing the effects of redesigning MOS when no external validation is possible. However, the ultimate test of utility will be their operational effectiveness.

One final observation concerns the data bases we studied, each of which suffered from at least one important limitation. The DOT and Project A data included the population of Army jobs, but the job descriptors would be inadequate for forming job families for most purposes other than research. The DOT data base contains descriptors of civilian jobs that were matched to Army MOS, and consequently, may not be accurate reflections of the Army jobs. In many cases the same descriptors applied to many MOS, e.g., electronics jobs.

The Project A data base consists of MOS groupings made by SMEs, which were used to form a distance matrix. These data do not provide individual variable scores for jobs and precluded some important analyses, i.e., the internal validity analysis based on the CCC index and the cross-validation. Although this data collection strategy requires less expenditure of project resources than a traditional job analysis study, it does not provide much information about the nature of the jobs.

The synthetic validity data base contained detailed data on the relevance, importance and difficulty of 96 tasks for different components of the job, but the sample of 21 jobs would be too small for most operational studies. None of the data bases we analyzed contained multiple ratings within jobs, which precluded the use of a better cross-validation procedure that leaves the job sample intact, and divides the sample of incumbents in half within each job.

More current and relevant job analysis information (on large samples, or the population of jobs), will be needed to support future research on, and changes in, Army jobs and job families. The job clustering studies reviewed in Chapter One, and the results we obtained, suggest strongly that special attention should be given to the linkage of job descriptors and clustering methods to the purpose for which jobs are being clustered. This includes considering the level of generality or specificity of the descriptor when selecting the job dimensions (e.g., tasks vs. behaviors vs. abilities), and developing the items. We obtained a relatively small number of clusters for the DOT jobs with all clustering procedures. Approximately twice as many were obtained in the Project A data, although 70% of the jobs were the same in the two data bases. Most of the difference was probably due to the type and level of specificity of job descriptors.

REFERENCES

Aldenderfer, M.S. & Blashfield, R.K. (1984). Cluster analysis. Newbury Park: Sage Publications.

Alley, W.E., Treat, B.R., & Black, D.E. (1988). Classification of Air Force jobs into aptitude clusters (AFHRL-TR-88-14). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.

Arvey, R.D., Maxwell, S.E., Gutenberg, R.L., & Camp, C. (1981). Detecting job differences: A Monte Carlo study. Personnel Psychology, 34, 709-730.

Arvey, R.D., Maxwell, S.E., & Mossholder, K.M. (1979). Even more ideas about methodologies for determining job differences and similarities. Personnel Psychology, 32, 529-538.

Arvey, R.D., & Mossholder, K.M. (1977). A proposed methodology for determining similarities and differences among jobs. Personnel Psychology, 30, 363-374.

Ballentine, R.D., Cunningham, J.W., & Wimpee, W.E. (1992). Air Force enlisted job clusters: An exploration in numerical job classification. Military Psychology, 4(2), 87-102.

Blashfield, R.K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. Psychological Bulletin, 83(3), 377-388.

Campbell, J.P. (1990). An overview of the Army selection and classification project (Project A). Personnel Psychology, 43, 231-239.

Cascio, W.F. (1987). Applied psychology in personnel management. (3rd ed.) Englewood Cliffs, NJ: Prentice-Hall, Inc.

Christal, R.E., & Weissmuller, J.J. (1988). Job-task inventory analysis. In S. Gael (ed.), The jobanalysis handbook for busineess, industry, and government: Volume II (pp. 1036-1050). New York: Wiley.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Colbert, G.A., & Taylor, L.R. (1978). Empirically derived job families as a foundation for the study of validity generalization. Study III. Generalization of selection test validity. Personnel Psychology, 31, 355-364.

Cornelius, E.T., III, Carron, T.J., & Collins, M.N. (1979). Job analysis models and job classification. Personnel Psychology, 32, 693-708.

Cornelius, E.T., III, Hakel, M.D., & Sackett, P.R. (1979). A methodological approach to job classification for performance appraisal purposes. Personnel Psychology, 32, 283-297.

Cornelius, E.T., III, Schmidt, F.L., & Carron, T.J. (1984). Job classification approaches and the implementation of validity generalization results. Personnel Psychology, 37, 247-259.

Cunningham, J.W., Wimpee, W.E. & Ballentine, R.D. (1990). Some general dimensions of work among U.S. Air Force enlisted occupations. Military Psychology, 2(1), 33-45.

Department of the Army (1986). Enlisted career management fields and military occupational specialties (Army Regulation 611-201). Washington, DC: Department of the Army.

Fraboni, M. & Cooper, D. (1989). Six Clustering algorithms applied to the WAIS-R: The problem of dissimilar cluster results. Journal of Clinical Psychology, 45(6), 932-935.

Garwood, M.K., Anderson, L.E., Greengart, B.J. (1991). Determining job groups: Application of hierarchical agglomerative cluster analysis in different job analysis situations. Personnel Psychology, 44(4), 743-762.

Hanser, L.M., Mendel, R.M., & Wolins, L. (1979). Three flies in the ointment: A reply to Arvey and Mossholder. Personnel Psychology, 32, 511-516.

Harris, D.A., McCloy, R.A., Dempsey, J.R., DiFazio, A.S., & Hogan, P. (1993). Personnel enlistment testing, job performance and cost: A cost-effectiveness analysis. (Final Report FR-PRD-93-35 under Contract NO MDA903-91-C-0242.) Alexandria, VA: U.S. Army Research Institute for the Behavioral Sciences.

Harris, D.A., McCloy, R.A., Dempsey, J.R., Roth, C., Sackett P.R., & Hedges, L.V. (1991). Determining the relationship between recruit characteristics and job performance: a methodology and a model. (Final Report 90-17.) Washington, D.C.: Office of the Assistant Secretary of Defense, Force Management and Personnel.

Harvey, R.J. (1986). Quantitative approaches to job classification: A review and critique. Personnel Psychology, 39, 267-289.

Harvey, R.J. (1991). Job analysis. In M.D. Dunnette & L.M. Hough (Eds.) Handbook of Industrial and Organizational Psychology (pp. 165-207). Palo Alto: Consulting Psychologists Press, Inc.

Headquarters, Department of the Army. (1992). Enlisted career management fields and military occupational specialties (AR 611-201 Revision). In Update 12-4: Military occupational classification structure. Washington, DC: Author.

Hoffman, R.G. (1987). Clustering Army military occupational specialties for Project A: Phase II (HumRRO Interim Report IR-PRD-87-22). Alexandria, VA: Human Resources Research Organization.

Hoffman, R.G., Fotouhi, C.H., Campshure, D.A., & Chia, W.J. (1991). Analysis of the Army Task Questionnaire. In L.L. Wise, N.G. Peterson, R.G. Hoffman, J.P. Campbell, & J.M. Arabian (Eds.), Army synthetic validity project: Report of Phase III results, volume I (ARI Technical Report 922). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 355)

Hubert, L.J., & Arabie, P. (1985). Comparing partitions. Journal of Classification, 2, 193-218.

Jain, A.K, & Dubes, R.C. (1988). Algorithms for clustering data. Engelwood Cliffs, NJ: Prentice Hall.

Johnson, C.D., Zeidner, J., & Leaman, J.A. (1992). Improving classification efficiency by restructuring Army job families (ARI Technical Report 947). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences. (AD A250 139)

Lancaster, A.R. (1984). U.S. Department of Defense initiatives to enhance high school career exploration: An overview. Paper presented at the 26th Annual Conference of the Military Testing Association, Munich, Federal Republic of Germany.

Laurence, J.H., & Hoffman, R.G. (nd). A description and evaluation of selection and classification models. Alexandria, VA: Human Resources Research Organization.

Lee, J.A., & Mendoza, J.L. (1981). A comparison of techniques which test for job differences. Personnel Psychology, 34, 731-748.

Lissitz, R.W., Mendoza, J.L., Huberty, C.J., & Markos, H.V. (1979). Some further ideas on a methodology for determining job similarities/ differences. Personnel Psychology, 32, 517-528.

Lord, F.M. & Novick, M.R. (1968). Statistical theories of mental test scores. Menlo Park, CA: Addison-Wesley.

MacQueen, J.B. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.

Maier, M.H., & Fuchs, E.F. (1972). Development and evaluation of a new ACB and aptitude area system (Technical Note 239). Arlington, VA: US Army Behavioral Science Research Laboratory. (NTIS No. Ad-703 134).

Maier, M.H. & Grafton, F.C. (1981). Aptitude composites for ASVAB 8, 9, and 10 (Research Report 1308.) Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

McCormick, E.J. (1979). Job analysis: Methods and applications. New York: AMACOM.

McIntyre, R.M., & Blashfield, R.K. (1980). A nearest-centroid technique for evaluating the minimum-variance clustering procedure. Multivariate Behavioral Research, 15, 225-238.

McIntyre, R.M., & Farr, J.L. (1979). Comment on Arvey and Mossholder's "A proposed methodology for determining similarities and differences among jobs". Personnel Psychology, 32, 507-510.

Milligan, G.W. (1979). Ultrametric hierarchical clustering algorithms. Psychometrika, 44, 343-346.

Milligan, G.W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika, 45, 325-342.

Milligan, G.W. (1981a). A Monte-Carlo study of 30 internal criterion measures for cluster-analysis. Psychometrika, 46, 187-191.

Milligan, G.W. (1981b). A review of Monte Carlo tests of cluster analysis. Multivariate Behavioral Research, 16, 370-407.

Milligan, G.W. (1985). An algorithm for generating artificial test clusters. Psychometrika, 50, 123-127.

Milligan, G.W., & Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika, 50, 159-179.

Milligan, G.W. & Cooper, M.C. (1987). Methodology review: clustering methods. Applied Psychological Measurement, 11(4), 329-354.

Milligan, G.W., & Isaac, P.D. (1980). The validation of four ultrametric clustering algorithms . Pattern Recognition, 12, 41-51.

Milligan, G.W., & Schilling, D.A. (1985). Asymptotic and finite-sample characteristics of four external criterion measures. Multivariate Behavioral Research 20, 97-109.

Milligan, G.W. & Sokol, L.M. (1980). A two-stage clustering algorithm with robust recovery characteristics. Educational and Psychological Measurement, 40, 755-759.

Milligan, G.W., Soon, S.C., & Sokol, L.M. (1983). The effect of cluster size, dimensionality, and the number of clusters on recovery of true cluster structure. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI 5, 40-47.

Mobley, W.H. & Ramsay, R.S. (1973). Hierarchical clustering on the basis of inter-job similarity as a tool in validity generalization. Personnel Psychology, 26, 213-225.

Office of the Assistant Secretary of Defense (Force Management and Personnel). (1987). Occupational conversion manual, enlisted/officer/civilian (DoD 1312.1-M). Arlington, VA: Defense Manpower Data Center.

Pearlman, K. (1980). Job families: A review and discussion of their implications for personnel selection. Psychological Bulletin, 87(1), 1-28.

Peterson, N.G., Owens-Kurtz, C.K., & Rosse, R.L. (1991). Formation of job performance prediction equations and evaluation of their validity (chapter 4). In L.L. Wise, N.G. Peterson, R.G. Hoffman, J.P. Campbell, & J.M. Arabian (Eds.), Army synthetic validity project: Report of Phase III results, volume I (ARI Technical Report 922). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A235 635)

Rand, W.M. (1971). Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association, 66, 846-850.

Reynolds, D.H., Barnes, J.D., Harris, D.A & Harris, J.H. (1992). Analysis and clustering of entry-level Navy ratings (Report No. FR-PRD-92-20). Alexandria, VA: Human Resources Research Organization.

Rosse, R. L., Borman, W.C., Campbell C.H., & Osborn, W.C. (October, 1983). Grouping Army occupational specialties by judged similarity. Paper presented to the Military Testing Association Convention, Gulf Shores, Alabama.

SAS Institute Inc. (1990). SAS/STAT User's Guide: Volume 1. Cary, NC: SAS Institute Inc.

Sackett, P.R. (1988). Exploring strategies for clustering military occupations. In B.F. Green, Jr. & Alexandra K. Wigdor (Eds.), Linking military enlistment standards to job performance: report of a workshop. Washington, D.C.: National Academy Press.

Sackett, P.R., Cornelius, E.T., III, & Carron, T.J. (1981). A comparison of global judgment vs. task oriented approaches to job classification. Personnel Psychology, 34, 791-804.

Sarle, W.S. (1983). Cubic clustering criterion. SAS Technical Report A-108. Cary, NC: SAS Institute, Inc.

Scheibler, D. & Schneider, W. (1985). Monte Carlo Tests of the accuracy of cluster analysis algorithms: A comparison of hierarchical and nonhierarchical methods. Multivariate Behavioral Research, 20, 283-304.

Schoenfeldt, L.F. (1985). Clustering military occupational specialties: A review and critique. Washington, DC: Paper submitted to Committee on the Performance of Military Personnel, National Research Council, National Academy of Sciences.

Sokal, R.R. (1988). Unsolved problems in numerical taxonomy. In H.H. Bock (Ed.), Classification and Related Methods of Data Analysis. North-Holland: Elsevier Science Publishers, B.V.

Statman, M.A. (1993). Improving the effectiveness of employment testing through classification: Alternative methods of developing test composites for optimal job assignment and vocational counseling. Unpublished doctoral dissertation, George Washington University.

Stutzman, T.M. (1983). Within classification job differences. Personnel Psychology, 36, 503-516.

Taylor, L.R. (1978). Empirically derived job families as a foundation for the study of validity generalization. Study I. The construction of job families based on the component and overall dimensions of the PAQ. Personnel Psychology, 31, 325-340.

Taylor, L.R., & Colbert, G.A. (1978). Empirically derived job families as a foundation for the study of validity generalization. Study II. The construction of job families based on the company-specific PAQ dimensions. Personnel Psychology, 31, 341-353.

U.S. Department of Labor, Employment and Training Administration (1977). Dictionary of Occupational Titles (4th ed.). Washington, D.C.: U.S. Government Printing Office.

Uniform guidelines on employee selection procedures. (1980). Washington, D.C.: The Bureau of National Affairs, Inc.

Wise, L.L., Peterson, N.G., Hoffman, R.G., Campbell, J.P., & Arabian, J.M. (Eds.). (1991). Army synthetic validity project: Report of Phase III Results, Volume 1 (ARI Technical Report 922). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences. (AD A231 437)

Wright, G.J. (1984). Crosscoding military and civilian occupational classification systems. Paper presented at the 26th Annual Conference of the Military Testing Association, Munich Federal Republic of Germany.

Zimmerman, R., Jacobs, R., & Farr, J. (1982). A comparison of the accuracy of four methods of clustering jobs. Applied Psychological Measurement, 6(3), 353-366.

Appendix A

MOS Titles, Aptitude Areas and Career Management Fields

MOS Titles, Aptitude Areas and Career Management Field for
317 MOS in DOT, Project A and Synthetic Validity Data Bases

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
00B	Diver	GM	General Maintenance	51	General Engineering
03C	Physical Activities Spec	ST	Skilled Technical	71	Administration
05B	Radio Operator	SC	Surveillance/C		
05C	Radio Teletype Operator	SC	Surveillance/C		
05D	EW/SIGINT Emitter ID/Locator	ST	Skilled Technical	98	Signs Int/Elect Warfare Oper
05G	Signal Security Specialist	ST	Skilled Technical	98	Signals Intel/Electronic Warfare Oper
05H	EW/SIGINT Intercept-IMC	ST	Skilled Technical	98	Signs Int/Elect Warfare Oper
05K	EW/SIGINT N-M Interceptor	CO	Combat	98	Signs Int/Elect Warfare Oper
11B	Infantryman	CO	Combat	11	Infantry
11C	Indirect Fire Infantryman	CO	Combat	11	Infantry
11H	Heavy Anti-Armor Wpns Inftryman	CO	Combat	11	Infantry
11M	Fighting Vehicle Infantryman	CO	Combat	11	Infantry
12B	Combat Engineer	CO	Combat	12	Combat Engineering
12C	Bridge Crewman	CO	Combat	12	Combat Engineering
12E	Atomic Demo Munitions Spec	CO	Combat	12	Combat Engineering
12F	Engineer Tracked Veh Crewman	CO	Combat	12	Combat Engineering
13B	Cannon Crewman	FA	Field Artillery	13	Field Artillery
13C	TACFIRE Operations Specialist	ST	Skilled Technical	13	Field Artillery
13E	Cannon Fire Direction Speclst	ST	Skilled Technical	13	Field Artillery
13F	Fire Support Specialist	FA	Field Artillery	13	Field Artillery
13M	MLRS Crewmember	OF	Operators/Food	13	Field Artillery
13R	Field Artlry Firefindr Rdr Op	SC	Surveillance/Communication	13	Field Artillery
15D	Lance Missile Crewmember	OF	Operators/Food	13	Field Artillery
15E	Pershing Missile Crewmember	OF	Operators/Food	13	Field Artillery
15J	MLRS/LANCE Ops/FireDir Spec	FA	Field Artillery	13	Field Artillery
16B	HERCULES Missile Crewmember	OF	Operators/Food	14	Air Defense Artillery
16C	HERCULES Fire Control Crewmbr	OF	Operators/Food	14	Air Defense Artillery
16D	HAWK Missile Crewmember	OF	Operators/Food	14	Air Defense Artillery
16E	HAWK Fire Control Crewmember	OF	Operators/Food	14	Air Defense Artillery
16F	Light Air Def Artillery Crewmbr	OF	Operators/Food	14	Air Defense Artillery
16H	ADA Opertns-Intelligence Assist	OF	Operators/Food	14	Air Defense Artillery
16J	Def Acq Radar Operator	OF	Operators/Food	14	Air Defense Artillery
16L	Sgt York Air Def Gun Crwmbr	OF	Operators/Food	14	Air Defense Artillery
16P	ADA Short Range Gunnery Crew	OF	Operators/Food	14	Air Defense Artillery
16R	ADA Short Range Gunry Crwmbr	OF	Operators/Food	14	Air Defense Artillery
16S	MANPADS Crewman	OF	Operators/Food	14	Air Defense Artillery
16T	Patriot Missile Crewmember	OF	Operators/Food	14	Air Defense
17B	Field Artlry Radar Crwmbr	SC	Surveillance/Communications	13	Field Artillery
17C	Field Artlry Target Acq Spec	SC	Surveillance/Communications	13	Field Artillery
17L	Aerial Sensor Specialist	ST	Skilled Technical	96	Military Intelligence
17M	Remote Sensor Specialist	SC	Surveillance/Communications	96	Military Intelligence
19D	Cavalry Scout	CO	Combat	19	Armor
19E	M48-M60 Armor Crewman	CO	Combat	19	Armor
19K	M1 ABRAMS Armor Crewman	CO	Combat	19	Armor
21G	PERSHING Elec Materiel Spec	EL	Electronics	13	Field Artillery
21L	PERSHING Electronics Repairer	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Mai
22L	NIKE Test Equipment Repairer	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Mai
22N	NIKEHER Missile Launcher Repair	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Mai

MOS Titles, Aptitude Areas and Career Management Field for
317 MOS in DOT, Project A and Synthetic Validity Data Bases (Cont.)

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
23N	NIKE Track Radar Repairer	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Maint
23U	NIKE Radar-Simulator Repairer	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Maint
24C	Improvd HAWK Firng Sect Mech	EL	Electronics	23	Air Defense System Maintenance
24E	Improvd HAWK Fire Contrl Mech	EL	Electronics	23	Air Defense System Maintenance
24G	Imp HAWK Inform CoorCentMech	EL	Electronics	23	Air Defense System Maintenance
24H	Improvd HAWK Fire Contrl Repr	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Maint
24J	Improved HAWK Pulse Radar Rep	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Maint
24K	ImpHAWK Cont-Wave Radar Repr	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Maint
24L	ImpHAWK Launch&Mech Sys Repr	EL	Electronics	27	Land Comb/Air Def Syst Dir/Gen Supp Maint
24M	VULCAN Systems Mechanic	EL	Electronics	23	Air Defense System Maintenance
24N	CHAPARRAL System Mechanic	EL	Electronics	23	Air Defense System Maintenance
24P	Defense Acq Radar Mechanic	EL	Electronics	23	Air Defense System Maintenance
24Q	NIKE-HERCULES Fire Contrl Mec	EL	Electronics	23	Air Defense System Maintenance
24T	PATRIOT Op & System Mechanic	MM	Mechanical Maintenance	23	Air Defense System Maintenance
24U	HERCULES Electronic Mechanic	EL	Electronics	23	Air Defense System Maintenance
24W	Sgt York Air Def Gun Syst Mec	EL	Electronics	23	Air Defense System Maintenance
25J		EL	Electronics		
25L	ADA Cmnd & Cntrl Syst Op/Repr	EL	Electronics	23	Air Defense System Maintenance
25S	Still Documentation Specialist	ST	Skilled Technicalcal	25	Visual Information
26B	Weapons Support Radar Repr	EL	Electronics	29	Signal Maintenance
26C	Tgt Acq/Surveillance Rdar Rep	EL	Electronics	29	Signal Maintenance
26D	Ground Cntrl Approch Rdar Rep	EL	Electronics	28	Avia Comm/Electronics Systems Maintenance
26E	Aerial Surv Sensor Repairer	EL	Electronics	33	Electronic Warfare/Intercept Systems Mai
26F	Aerial Photo-Activ Sensor Rep	EL	Electronics	33	Electronic Warfare/Intercept Systems Mai
26H	Air Defense Radar Repairer	EL	Electronics	23	Air Defense System Maintenance
26K	AerialEWarning/DefEquipRepair	SC	Surveillance/Communica	28	Avia Comm/Electronics Systems Maintenanc
26L	Tactical Microwave Syst Repr	EL	Electronics	29	Signal Maintenance
26M	Aerial Surveillance Radar Repr	EL	Electronics		
26N	Aerial Surveillance Infrared Repr	EL	Electronics		
26Q	Tact Satell/Microwave Syst Op	EL	Electronics	31	Signal Operations
26R	Strategic Microwave Syst Op	EL	Electronics	31	Signal Operations
26T	Radio/TV Systems Specialist	EL	Electronics	25	Visual Information
26V	Strategic Microwave Syst Repr	EL	Electronics	29	Signal Maintenance
26Y	SATCOM Equipment Repairer	EL	Electronics	29	Signal Maintenance
27B	LandCombat SystemTestSpecial	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27E	TOW/DRAGON Repairer	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27F	VULCAN Repairer	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27G	CHAPARRAL/REDEYE Repairer	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27H	HAWK Firing Section Reparirer	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27L	LANCE System Repairer	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27M	MLRS Repairer	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27N	Forwrd Area Alerting Rdar Rep	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27P	SgtYork Radar/Electron Repr	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
27Q	SgtYork Test Specialist	EL	Electronics	27	Land Comb/Air System Dir/Gen Supp Maint
29E	Radio Repairer	EL	Electronics	29	Signal Maintenance
31C	Single Channel Radio Operator	SC	Surveillance/Communica	31	Signal Operations
31D	Mobile Subscriber Equipment	EL	Electronics	31	Signal Operations
31E	Field Radio Repairer	EL	Electronics	29	Signal Maintenance

MOS Titles, Aptitutde Areas and Career Management Field for
317 MOS in DOT, Proejct A and Synthetic Validity Data Bases (Cont.)

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
31J	Teletypewriter Repairer	EL	Electronics	29	Signal Maintenance
31K	Combat Signaler	EL	Electronics	31	Signal Operations
31M	Multichannel Commo Equip Op	EL	Electronics	31	Signal Operations
31N	Tactical Circuit Controller	EL	Electronics	31	Signal Operations
31S	Field General Comsec Repairer	EL	Electronics	29	Signal Maintenance
31T	Field Systems Comsec Repairer	EL	Electronics	29	Signal Maintenance
31V	Tactical Commo Syst Op/Mech	EL	Electronics	31	Signal Operations
32D	Station Technical Controller	EL	Electronics	31	Signal Operations
32F	Fixed Ciphony Repairer	EL	Electronics	29	Signal Maintenance
32G	Fixed Crypto Equip Repairer	EL	Electronics	29	Signal Maintenance
32H	Fixed Station Radio Repairer	EL	Electronics	29	Signal Maintenance
33P	Ele War/Int StrRecv Subt Repr	ST	Skilled Technical	33	Electronic Warfare/Intercept Systems Mai
33Q	Ele War/Int StratProcess&Stor	ST	Skilled Technical	33	Electronic Warfare/Intercept Systems Mai
33R	EW/Intrcpt AvSystems Repairer	ST	Skilled Technical	33	Electronic Warfare/Intercept Systems Mai
33S	EW/Intercept Sys Repr	ST	Skilled Technical	33	Electronic Warfare/Intercept Systems Mai
33T	Ele War/Int Tact Systems Rep	ST	Skilled Technical	33	Electronic Warfare/Intercept Systems Mai
34B	Punch Card Machine Operator	EL	Electronics	74	Record Information Operations
34C	Decen Auto Serv Supp Systm	EL	Electronics	74	Record Information Operations
34E	NCR 500 Computer Repairer	EL	Electronics	74	Record Information Operations
34F	Digt Subsc Message Switch Equip	EL	Electronics	74	Record Information Operations
34H	Auto Digt Message Switch Equip	EL	Electronics	74	Record Information Operations
34L	Field Artlry Digital Sys Rep	EL	Electronics	74	Record Information Operations
34T	Tact Computer Systems Repr	EL	Electronics	74	Record Information Operations
34Y	Field Artlry TactFire Repair	EL	Electronics	74	Record Information Operations
35B	Electronics Instrument Repr	EL	Electronics		
35C	Automatic Test Equip Repairer	EL	Electronics	74	Record Information Operations
35E	Special Elec Devices Repairer	EL	Electronics	29	Signal Maintenance
35F	Nuclear Weapons Electronics Specialist	EL	Electronics		
35G	Biomedical Equipment Spec	EL	Electronics	91	Medical
35H	Calibration Specialist	EL	Electronics	35	Electronic Maintenance
35K	Avionic Mechanic	EL	Electronics	28	Avia Comm/Electronics Systems Maintenance
35L	Avionic Commo Equip Repairer	EL	Electronics	28	Avia Comm/Electronics Systems Maintenance
35M	Avionic Nav/FlightContEq Repr	EL	Electronics	28	Avia Comm/Electronics Systems Maintenance
35R	Avionic Special Equip Repr	EL	Electronics	28	Avia Comm/Electronics Systems Maintenance
36C	Wire System Instll/Operator	EL	Electronics	31	Signal Operations
36D	Antenna Installer Specialist	EL	Electronics		
36E	Cable Splicer	EL	Electronics		
36H	Dial/Manual Centrl Office Rep	EL	Electronics	29	Signal Maintenance
36K	Tactical Wire Operations Specialist	EL	Electronics		
36L	Trans ElectSwitchSys Rep	EL	Electronics	31	Signal Operations
36M	Wire Systems Operator	EL	Electronics	31	Signal Operations
39G	Auto Comm Comp Syst Repairer	EL	Electronics	29	Signal Maintenance
41B	Topographic Instr Rep Spec	GM	General Maintenance	81	Topographic Engineering
41C	Fire Contrl Instru Rep Spec	GM	General Maintenance	63	Mechanical Maintenance
41E	Audio/Visual Equip Repairer	EL	Electronics	25	Visual Information
41G	Aerial Surveillance Photo Equip Repr	EL	Electronics		
41J	Office Machine Repairer	GM	General Maintenance	63	Mechanical Maintenance
42C	Orthotic Specialist	GM	General Maintenance	91	Medical

MOS Titles, Aptitutde Areas and Career Management Field for
317 MOS in DOT, Proejct A and Synthetic Validity Data Bases (Cont.)

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
42D	Dental Laboratory Specialist	GM	General Maintenance	91	Medical
42E	Optical Laboratory Spec	GM	General Maintenance	91	Medical
43E	Parachute Rigger	GM	General Maintenance	76	Supply and Services
43M	Fabric Repair Specialist	GM	General Maintenance	76	Supply and Services
44B	Metalworker	GM	General Maintenance	63	Mechanical Maintenance
44E	Machinist	GM	General Maintenance	63	Mechanical Maintenance
45B	Small Arms Repairer	GM	General Maintenance	63	Mechanical Maintenance
45D	SP Field Artlry Turret Mech	GM	General Maintenance	63	Mechanical Maintenance
45E	M1 ABRAMS Tank Turret Mech	GM	General Maintenance	63	Mechanical Maintenance
45G	Fire Control System Repairer	EL	Electronics	63	Mechanical Maintenance
45K	Tank Turret Repairer	GM	General Maintenance	63	Mechanical Maintenance
45L	Artillery Repairer	GM	General Maintenance	63	Mechanical Maintenance
45N	M60A1/A3 Tank Turret Mech	MM	Mechanical Maintenanc	63	Mechanical Maintenance
45T	BFVS Turret Mechanic	GM	General Maintenance	63	Mechanical Maintenance
46N	PERSHING ElecMechcal Repairer	EL	Electronics	63	Mechanical Maintenance
46Q	Journalist	GT	General Technical	46	Public Affairs
51B	Carpentry/Masonry Specialist	GM	General Maintenance	51	General Engineering
51C	Structures Specialist	GM	General Maintenance	51	General Engineering
51G	Materials Quality Specialist	GM	General Maintenance	51	General Engineering
51K	Plumber	GM	General Maintenance	51	General Engineering
51M	Firefighter	GM	General Maintenance	51	General Engineering
51N	Water Treatment Specialist	GM	General Maintenance	77	Petroleum and Water
51R	Interior Electrician	GM	General Maintenance	51	General Engineering
52C	Utilities Equipment Repairer	GM	General Maintenance	63	Mechanical Maintenance
52D	Power Generator Equip Repr	GM	General Maintenance	63	Mechanical Maintenance
52E	Prime Power Prod Specialist	ST	Skilled Technical	51	General Engineering
52F	Turbine Engine Generator Repr	GM	General Maintenance	63	Mechanical Maintenance
52G	Transmisson & Distbution Spec	EL	Electronics	51	General Engineering
53B	Industrial Gas Prod Specialist	GM	General Maintenance	55	Ammunition
54B	Chemical Operations Specialist	ST	Skilled Technical	54	Chemical
54C	Smoke Operation Specialist	GM	General Maintenance	54	Chemical
54E	NBC Specialist	ST	Skilled Technical	54	Chemical
55B	Ammunition Specialist	GM	General Maintenance	55	Ammunition
55D	Explsve Ordnance Disposl Spec	GM	General Maintenance	55	Ammunition
55G	Nuclear Weapons Maint Spec	GM	General Maintenance	55	Ammunition
55R	Ammo Stock Control&Acct Spec	ST	Skilled Technical	55	Ammunition
57E	Laundry & Bath Specialist	GM	General Maintenance	76	Supply and Services
57F	Graves Registration Spec	GM	General Maintenance	76	Supply and Services
57H	Cargo Specialist	GM	General Maintenance	88	Transportation
61B	Watercraft Operator	MM	Mechanical Maintenance	88	Transportation
61C	Watercraft Engineer	MM	Mechanical Maintenance	88	Transportation
61F	Marine Hull Repr	GM	General Maintenance		
62B	Construction Equipment Repr	MM	Mechanical Maintenance	63	Mechanical Maintenance
62E	Hvy Construction Equip Op	GM	General Maintenance	51	General Engineering
62F	Crane Operator	GM	General Maintenance	51	General Engineering
62G	Quarrying Specialist	GM	General Maintenance	51	General Engineering
62H	Concrete & Asphalt Equip Op	GM	General Maintenance	51	General Engineering
62J	General Construc Equip Op	GM	General Maintenance	51	General Engineering

MOS Titles, Aptitutde Areas and Career Management Field for
317 MOS in DOT, Proejct A and Synthetic Validity Data Bases (Cont.)

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
63B	Light Wheel Vehicle Mechanic	MM	Mechanical Maintenance	63	Mechanical Maintenan
63D	SP Field Artilry System Mech	MM	Mechanical Maintenance	63	Mechanical Maintenan
63E	M1 Abrams Tank System Mech	MM	Mechanical Maintenance	63	Mechanical Maintenan
63G	Fuel & Elec System Repairer	MM	Mechanical Maintenance	63	Mechanical Maintenan
63H	Track Vehicle Repairer	MM	Mechanical Maintenance	63	Mechanical Maintenan
63J	Quart&Chem Equipment Repairer	MM	Mechanical Maintenance	63	Mechanical Maintenan
63N	M60A1/A3 Tank System Mechanic	MM	Mechanical Maintenance	63	Mechanical Maintenan
63S	Heavy Wheel Vehicle Mechanic	MM	Mechanical Maintenance	63	Mechanical Maintenan
63T	Bradley System Mechanic	MM	Mechanical Maintenance	63	Mechanical Maintenan
63W	Wheel Vehicle Repairer	MM	Mechanical Maintenance	63	Mechanical Maintenan
63Y	Track Vehicle Mechanic	MM	Mechanical Maintenance	63	Mechanical Maintenan
64C	Motor Transport Operator	OF	Operators/Food	88	Transportation
65B	Locomotive Repairer	MM	Mechanical Maintenance	88	Transportation
65D	Railway Car Repairer	MM	Mechanical Maintenance	88	Transportation
65E	Airbrake Repairer	MM	Mechanical Maintenance	88	Transportation
65H	Locomotive Operator	MM	Mechanical Maintenance	88	Transportation
65J	Train Crewmemeber	MM	Mechanical Maintenance	88	Transportation
67G	Utility Airplane Repairer	MM	Mechancial Maintenance	67	Aircraft Maintenance
67H	Observation Airplane Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
67N	Utility Helicopter Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
67R	AH-64 Attack Helicoptr Repr	MM	Mechanical Maintenance	67	Aircraft Maintenance
67S	Scout Helicopter Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
67T	Tact Transp Helicoptr Repr	MM	Mechanical Maintenance	67	Aircraft Maintenance
67U	Medium Helicopter Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
67V	Observ/Scout Helicoptr Repr	MM	Mechanical Maintenance	67	Aircraft Maintenance
67X	Heavy Lift Helicopter Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
67Y	AH-1 Attack Helicoptr Repr	MM	Mechanical Maintenance	67	Aircraft Maintenance
68B	Aircraft Powerplant Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
68D	Aircraft Powertrain Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
68F	Aircraft Electrician	MM	Mechanical Maintenance	67	Aircraft Maintenance
68G	Aircraft Structural Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
68H	Aircraft Pneudraulic Repairer	MM	Mechanical Maintenance	67	Aircraft Maintenance
68J	Aircraft Fire Control Repr	MM	Mechanical Maintenance	67	Aircraft Maintenance
68M	Aircraft Weapon Systems Repr	GM	General Maintenance	67	Aircraft Maintenance
68N	Avionic Mechanic	EL	Electronics	67	Aircraft Maintenance
71C	Exec Administrative Assistant	CL	Clerical	71	Administration
71D	Legal Clerk Specialist	CL	Clerical	71	Administration
71E	Court Reporter	CL	Clerical	71	Administration
71G	Patient Admin Specialist	CL	Clerical	91	Medical
71L	Administrative Specialist	CL	Clerical	71	Administration
71M	Chapel Activities Specialist	CL	Clerical	71	Administration
71N	Traffic Management Coordntor	CL	Clerical	88	Transportation
71P	Flight Operations Coordinator	ST	Skilled Technical		
71Q	Journalist	ST	Skilled Technical	46	Public Affairs
71R	Broadcast Journalist	ST	Skilled Technical	46	Public Affairs
72E	Combat Telecomm Ctr Operator	SC	Surveillance/Communica	31	Signal Operations
72G	Auto Data Telecomm Ctr Oprtor	SC	Surveillance/Communica	31	Signal Operations
72H	Central Office Operations Op	CL	Clerical	31	Signal Operations

MOS Titles, Aptitutde Areas and Career Management Field for

317 MOS in DOT, Proejct A and Synthetic Validity Data Bases

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
73C	Finance Specialist	CL	Clerical	71	Administration
73D	Accounting Specialist	ST	Skilled Technical	71	Administration
74B	Card and Tape Writer	CL	Clerical	74	Record Information Operations
74C	Rec Telecom Cen Opr	EL	Electronics	74	Record Information Operations
74D	Computer/Machine Operator	ST	Skilled Technical	74	Record Information Operations
74F	Programmer Analyst	ST	Skilled Technical	74	Record Information Operations
75B	Personnel Admin Specialist	CL	Clerical	71	Administration
75C	Personnel Management Spec	CL	Clerical	71	Administration
75D	Personnel Records Specialist	CL	Clerical	71	Administration
75E	Personnel Action Specialist	CL	Clerical	71	Administration
75F	Personnel Info Mangmt Spec	CL	Clerical	71	Administration
76C	Equip Records & Parts Spec	CL	Clerical	76	Supply and Services
76J	Medical Supply Specialist	CL	Clerical	91	Medical
76P	Materl Centrl & Acctng Spec	CL	Clerical	76	Supply and Services
76V	Mat Storage & Handlng Spec	CL	Clerical	76	Supply and Services
76W	Petroleum Supply Specialist	CL	Clerical	77	Petroleum and Water
76X	Subsistence Supply Specialist	CL	Clerical	76	Supply and Services
76Y	Unit Supply Specialist	CL	Clerical	76	Supply and Services
81B	Technical Drafting Specialist	ST	Skilled Technical	51	General Engineering
81C	Cartographer	ST	Skilled Technical	81	Topographic Engineering
81E	Illustrator	ST	Skilled Technical	25	Visual Information
81Q	Terrain Analyst	ST	Skilled Technical	81	Topographic Engineering
82B	Construction Surveyor	ST	Skilled Technical	51	General Engineering
82C	Field Artillery Surveyor	ST	Skilled Technical	13	Field Artillery
82D	Topographic Surveyor	ST	Skilled Technical	81	Topographic Engineering
83E	Photo & Layout Specialist	ST	Skilled Technical	81	Topographic Engineering
83F	Photolithographer	ST	Skilled Technical	81	Topographic Engineering
84B	Still Photographic Specialist	ST	Skilled Technical	25	Visual Information
84C	Motion Picture Specialist	ST	Skilled Technical	25	Visual Information
84F	Audio/TV Specialist	ST	Skilled Technical	25	Visual Information
88L	Watercraft Engineer	MM	Mechanical Maintenance	88	Transportation
88M	Motor Transport Operator	OF	Operations Food	88	Transportation
91A	Medical Specialist	ST	Skilled Technical	91	Medical
91C	Practical Nurse	ST	Skilled Technical	91	Medical
91D	Operating Room Specialist	ST	Skilled Technical	91	Medical
91E	Dental Specialist	ST	Skilled Technical	91	Medical
91F	Psychiatric Specialist	ST	Skilled Technical	91	Medical
91G	Behavioral Science Specialist	ST	Skilled Technical	91	Medical
91H	Orthopedic Specialist	ST	Skilled Technical	91	Medical
91J	Physical Therapy Specialist	ST	Skilled Technical	91	Medical
91L	Occupational Therapy Spec	ST	Skilled Technical	91	Medical
91N	Cardiac Specialist	ST	Skilled Technical	91	Medical
91P	X-Ray Specialist	ST	Skilled Technical	91	Medical
91Q	Pharmacy Specialist	ST	Skilled Technical	91	Medical
91R	Veterinary Food Inspec Spec	ST	Skilled Technical	91	Medical
91S	Environmental Health Spec	ST	Skilled Technical	91	Medical
91T	Animal Care Specialist	ST	Skilled Technical	91	Medical
91U	Ear, Nose & Throat Specialist	ST	Skilled Technical	91	Medical

MOS Titles, Aptitutde Areas and Career Management Field for
317 MOS in DOT, Proejct A and Synthetic Validity Data Bases

MOS	MOSTITLE	AAID	AATITLE	CMFNO	CMFTITLE
91V	Respiratory Specialist	ST	Skilled Technical	91	Medical
91Y	Eye Specialist	ST	Skilled Technical	91	Medical
92B	Medical Laboratory Specialist	ST	Skilled Technical	91	Medical
92C	Petroleum Laboratory Spec	ST	Skilled Technical	77	Petroleum and Water
92D	Chemical Laboratory Spec	ST	Skilled Technical	54	Chemical
93C	Air Traffic Control	ST	Skilled Technical	93	Aviation Operations
93E	Meteorological Observer	ST	Skilled Technical	93	Aviation Operations
93F	Field Artlry Meteorlogic Spec	EL	Electronics	13	Field Artillery
93H	Air Traffic Control Tower Op	ST	Skilled Technical	93	Aviation Operations
93J	Air Traff Cntrl Radar Contlr	ST	Skilled Technical	93	Aviation Operations
93P	Flight Operations Coordinator	ST	Skilled Technical	93	Aviation Operations
94B	Food Service Specialist	OF	Operators/Food	94	Food Services
94F	Hospital Food Service Spec	OF	Operators/Food	91	Medical
95B	Military Police	ST	Skilled Technical	95	Military Police
95C	Correctional Specialist	ST	Skilled Technical	95	Military Police
95D	Special Agent	ST	Skilled Technical	95	Military Police
96B	Intelligence Analyst	ST	Skilled Technical	96	Military Intelligence
96C	Interrogator	ST	Skilled Technical	96	Military Intelligence
96D	Imagery Analyst	ST	Skilled Technical	96	Military Intelligence
96F	Psychological Opertns Spec	ST	Skilled Technical	96	Military Intelligence
96H	Aerial Intell Spec	ST	Skilled Technical	96	Military Intelligence
96R	Ground Surv Systems Operator	SC	Surveillance/Communica	96	Military Intelligence
97B	Counterintelligence Agents	ST	Skilled Technical	96	Military Intelligence
97E	Interrogator	ST	Skilled Technical	96	Military Intelligence
97G	Signal Security Specialist	ST	Skilled Technical	96	Military Intelligence
98C	EW/SIGINT Analyst	ST	Skilled Technical	98	Signals Intel/Electronic Warfare Oper
98D	Emitter Locator/Identifier	ST	Skilled Technical	98	Signals Intel/Electronic Warfare Oper
98G	EW/SIGINT Voice Interceptor	ST	Skilled Technical	98	Signals Intel/Electronic Warfare Oper
98J	EW/SIGINT NoncommoIntercept	ST	Skilled Technical	98	Signals Intel/Electronic Warfare Oper

Appendix B

Ward Hierarchical Cluster Analysis of DOT Data

- Plot of Cubic Cluster Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
PLOT OF CUBIC CLUSTERING CRITERION BY NUMBER OF CLUSTERS

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.

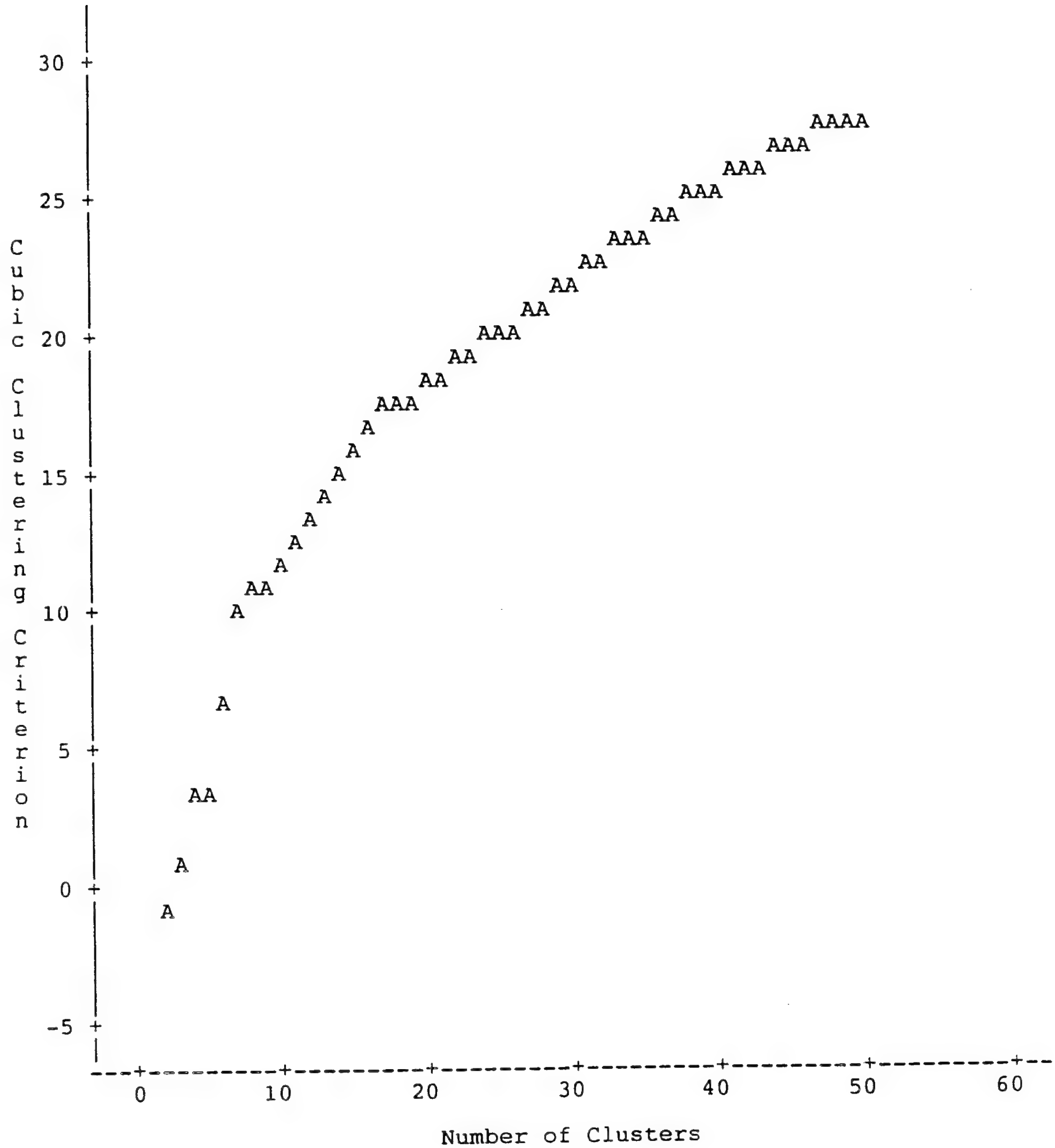


Figure 1. CCC Values by Number of Clusters

PLOT OF GAMMA VALUES BY NUMBER OF CLUSTERS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

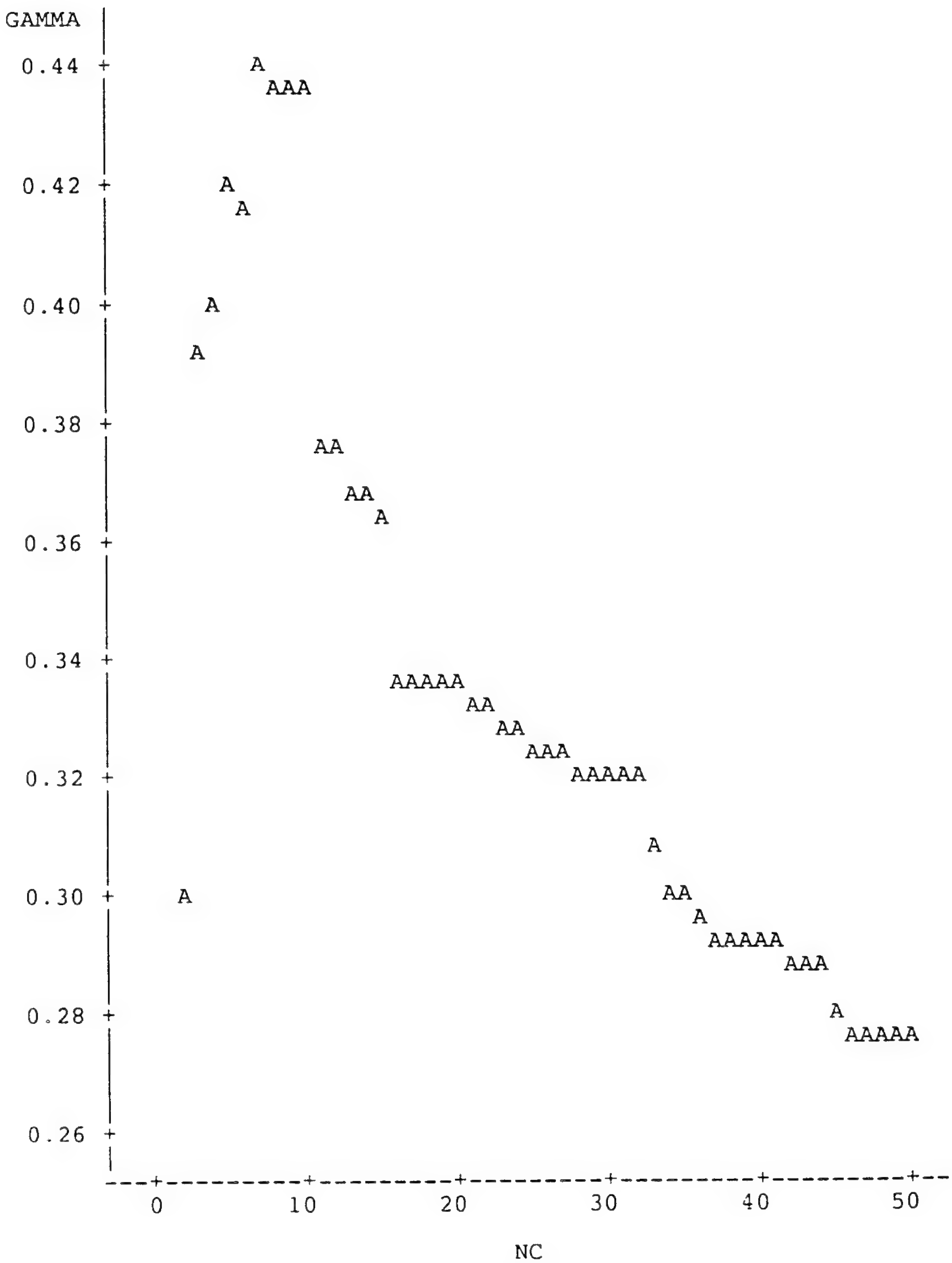


Figure 2. Gamma Values by Number of Clusters

Table 1.
WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
SEVEN CLUSTER SOLUTION

------------------------------ CLSNUM=1 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
1	71G	Patient Admin Specialist	0.39397
2	72E	Combat Telecomm Ctr Operator	0.44179
3	75B	Personnel Admin Specialist	0.58824
4	75D	Personnel Records Specialist	0.58824
5	75E	Personnel Action Specialist	0.58824
6	75F	Personnel Info Mangmt Spec	0.58824
7	91H	Orthopedic Specialist	0.61599
8	76C	Equip Records & Parts Spec	0.67646
9	91E	Dental Specialist	0.70011
10	76P	Materl Centrl & Acctng Spec	0.87734
11	05D	EW/SIGINT Emitter ID/Locator	0.88312
12	05G	Signal Security Specialist	0.88312
13	05H	EW/SIGINT Intercept-IMC	0.88312
14	05K	EW/SIGINT N-M Interceptor	0.88312
15	72G	Auto Data Telecomm Ctr Oprtor	0.93878
16	74D	Computer/Machine Operator	0.93878
17	05C	Radio Teletype Operator	0.94256
18	91L	Occupational Therapy Spec	0.95307
19	91U	Ear, Nose & Throat Specialist	0.95975
20	71L	Administrative Specialist	0.97651
21	73D	Accounting Specialist	1.00168
22	65E	Airbrake Repairer	1.12011
23	35H	Calibration Specialist	1.17073
24	91T	Animal Care Specialist	1.23784
25	94B	Food Service Specialist	1.25396
26	94F	Hospital Food Service Spec	1.25396
27	74B	Card and Tape Writer	1.26327
28	76Y	Unit Supply Specialist	1.28821
29	91Y	Eye Specialist	1.33538
30	91C	Practical Nurse	1.36440
31	91F	Psychiatric Specialist	1.36851
32	73C	Finance Specialist	1.41379
33	76J	Medical Supply Specialist	1.41445
34	76V	Mat Storage & Handlng Spec	1.41445
35	76X	Subsistence Supply Specialist	1.41445
36	31N	Tactical Circuit Controller	1.46615
37	36M	Wire Systems Operator	1.46615
38	43E	Parachute Rigger	1.49691
39	65H	Locomotive Operator	1.52121
40	36K	Tactical Wire Operations Specialist	1.55858
41	43M	Fabric Repair Specialist	1.67314
42	91R	Veterinary Food Inspec Spec	1.77308
43	91S	Environmental Health Spec	1.77308
44	53B	Industrial Gas Prod Specialist	2.07207
45	57F	Graves Registration Spec	2.32522
46	57H	Cargo Specialist	2.48974

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
SEVEN CLUSTER SOLUTION

------------------------------ CLSNUM=2 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
47	19E	M48-M60 Armor Crewman	0.41591
48	19K	M1 ABRAMS Armor Crewman	0.41591
49	13B	Cannon Crewman	0.60199
50	62H	Concrete & Asphalt Equip Op	0.63753
51	54C	Smoke Operation Specialist	0.67362
52	62J	General Construc Equip Op	0.68339
53	64C	Motor Transport Operator	0.80597
54	11C	Indirect Fire Infantryman	0.85503
55	61B	Watercraft Operator	0.92262
56	11B	Infantryman	1.09223
57	65J	Train Crewmemeber	1.11903
58	22N	NIKEHER MissileLauncherRepair	1.13659
59	24L	ImpHAWK Launch&Mech Sys Repr	1.13659
60	27H	HAWK Firing Section Reparirer	1.13659
61	46N	PERSHING ElecMechcal Repairer	1.13659
62	12B	Combat Engineer	1.42600
63	16F	Light Air Def Artillery Crewmbr	1.55203
64	16L	Sgt York Air Def Gun Crwmbr	1.55203
65	16R	ADA Short Range Gunry Crwmbr	1.55203
66	55B	Ammunition Specialist	1.60722
67	55R	Ammo Stock Control&Acct Spec	1.60722
68	11H	Heavy Anti-Armor Wpns Inftryman	1.83097
69	11M	Fighting Vehicle Infantryman	1.83097
70	12C	Bridge Crewman	1.89756
71	51N	Water Treatment Specialist	1.93592
72	57E	Laundry & Bath Specialist	1.97951
73	54E	NBC Specialist	2.45637
74	17C	Field Artlry Target Acq Spec	2.47599

------------------------------ CLSNUM=3 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
75	93H	Air Traffic Control Tower Op	0.68114
76	93J	Air Traff Cntrl Radar Contlr	0.68114
77	13E	Cannon Fire Direction Speclst	0.74084
78	15J	MLRS/LANCE Ops/FireDir Spec	0.74084
79	91N	Cardiac Specialist	0.84930
80	71C	Exec Administrative Assistant	0.97723
81	42C	Orthotic Specialist	1.05111
82	96D	Imagery Analyst	1.12369
83	95B	Military Police	1.19352
84	17L	Aerial Sensor Specialist	1.19411
85	96H	Aerial Intell Spec	1.19411
86	91P	X-Ray Specialist	1.23100
87	84B	Still Photographic Specialist	1.26883
88	84C	Motion Picture Specialist	1.26883
89	84F	Audio/TV Specialist	1.26883
90	17M	Remote Sensor Specialist	1.36572
91	13F	Fire Support Specialist	1.49085
92	91J	Physical Therapy Specialist	1.50076
93	81C	Cartographer	1.54412

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
SEVEN CLUSTER SOLUTION

------------------------------ CLSNUM=3 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
94	71E	Court Reporter	1.55296
95	82B	Construction Surveyor	1.58926
96	91A	Medical Specialist	1.61785
97	95C	Correctional Specialist	1.72784
98	19D	Cavalry Scout	2.10109
99	03C	Physical Activities Spec	2.22442
100	81E	Illustrator	2.39148

------------------------------ CLSNUM=4 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
101	21L	PERSHING Electronics Repairer	0.38655
102	22L	NIKE Test Equipment Repairer	0.38655
103	23N	NIKE Track Radar Repairer	0.38655
104	23U	NIKE Radar-Simulator Repairer	0.38655
105	24C	Improvd HAWK Firng Sect Mech	0.38655
106	24E	Improvd HAWK Fire Contrl Mech	0.38655
107	24G	Imp HAWK Inform CoorCentMech	0.38655
108	24H	Improvd HAWK Fire Contrl Repr	0.38655
109	24J	Improved HAWK Pulse Radar Rep	0.38655
110	24K	ImpHAWK Cont-Wave Radar Repr	0.38655
111	24P	Defense Acq Radar Mechanic	0.38655
112	24Q	NIKE-HERCULES Fire Contrl Mec	0.38655
113	24U	HERCULES Electronic Mechanic	0.38655
114	24W	Sgt York Air Def Gun Syst Mec	0.38655
115	25J		0.38655
116	26B	Weapons Support Radar Repr	0.38655
117	26D	Ground Cntrl Approch Rdar Rep	0.38655
118	26E	Aerial Surv Sensor Repairer	0.38655
119	26H	Air Defense Radar Repairer	0.38655
120	26M	Aerial Surveillance Radar Repr	0.38655
121	26Y	SATCOM Equipment Repairer	0.38655
122	27B	LandCombat SystemTestSpecial	0.38655
123	27E	TOW/DRAGON Repairer	0.38655
124	27F	VULCAN Repairer	0.38655
125	27G	CHAPARRAL/REDEYE Repairer	0.38655
126	27L	LANCE System Repairer	0.38655
127	27M	MLRS Repairer	0.38655
128	27N	Forwrd Area Alerting Rdar Rep	0.38655
129	27P	SgtYork Radar/Electron Repr	0.38655
130	27Q	SgtYork Test Specialist	0.38655
131	32F	Fixed Ciphony Repairer	0.38655
132	32G	Fixed Crypto Equip Repairer	0.38655
133	33S	EW/Intercept Sys Repr	0.38655
134	34C	Decen Auto Serv Supp Systm	0.38655
135	34E	NCR 500 Computer Repairer	0.38655
136	34F	Digt Subsc Message Switch Equip	0.38655
137	34H	Auto Digt Message Switch Equip	0.38655
138	34Y	Field Artlry TactFire Repair	0.38655
139	35B	Electronics Instrument Repr	0.38655

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
SEVEN CLUSTER SOLUTION

------------------------------- CLSNUM=4 -------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
140	35C	Automatic Test Equip Repairer	0.38655
141	35F	Nuclear Weapons Electronics Specialist	0.38655
142	45G	Fire Control System Repairer	0.38655
143	41E	Audio/Visual Equip Repairer	0.50756
144	81B	Technical Drafting Specialist	0.51325
145	42D	Dental Laboratory Specialist	0.62258
146	34B	Punch Card Machine Operator	0.62267
147	41J	Office Machine Repairer	0.62828
148	26F	Aerial Photo-Activ Sensor Rep	0.66863
149	41G	Aerial Surveillance Photo Equip Repr	0.66863
150	35G	Biomedical Equipment Spec	0.68721
151	26N	Aerial Surveillance Infrared Repr	0.75073
152	31E	Field Radio Repairer	0.75073
153	31S	Field General Comsec Repairer	0.75073
154	31T	Field Systems Comsec Repairer	0.75073
155	32D	Station Technical Controller	0.75073
156	35E	Special Elec Devices Repairer	0.80939
157	41C	Fire Contrl Instru Rep Spec	0.80939
158	91Q	Pharmacy Specialist	0.83578
159	05B	Radio Operator	0.88184
160	26Q	Tact Satell/Microwave Syst Op	0.88184
161	26R	Strategic Microwave Syst Op	0.88184
162	31M	Multichannel Commo Equip Op	0.88184
163	31V	Tactical Commo Syst Op/Mech	0.88184
164	26T	Radio/TV Systems Specialist	0.97544
165	92D	Chemical Laboratory Spec	0.97613
166	91V	Respiratory Specialist	1.02500
167	42E	Optical Laboratory Spec	1.03111

------------------------------- CLSNUM=5 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
168	65B	Locomotive Repairer	0.13966
169	44B	Metalworker	0.23394
170	83F	Photolithographer	0.25333
171	61C	Watercraft Engineer	0.33162
172	31J	Teletypewriter Repairer	0.49581
173	36L	Trans ElectSwitchSys Rep	0.49581
174	83E	Photo & Layout Specialist	0.50341
175	41B	Topographic Instr Rep Spec	0.50464
176	26L	Tactical Microwave Syst Repr	0.53747
177	26V	Strategic Microwave Syst Repr	0.53747
178	32H	Fixed Station Radio Repairer	0.53747
179	68F	Aircraft Electrician	0.54942
180	44E	Machinist	0.57680
181	63B	Light Wheel Vehicle Mechanic	0.61703
182	63G	Fuel & Elec System Repairer	0.61703
183	63S	Heavy Wheel Vehicle Mechanic	0.61703
184	63W	Wheel Vehicle Repairer	0.61703
185	35K	Avionic Mechanic	0.62958

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
SEVEN CLUSTER SOLUTION

------------------------------- CLSNUM=5 -------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
186	35L	Avionic Commo Equip Repairer	0.62958
187	35M	Avionic Nav/FlightContEq Repr	0.62958
188	35R	Avionic Special Equip Repr	0.62958
189	36C	Wire System Instll/Operator	0.67395
190	52D	Power Generator Equip Repr	0.67975
191	67G	Utility Airplane Repairer	0.68459
192	67H	Observation Airplane Repairer	0.68459
193	67N	Utility Helicopter Repairer	0.68459
194	67T	Tact Transp Helicoptr Repr	0.68459
195	67U	Medium Helicopter Repairer	0.68459
196	67V	Observ/Scout Helicoptr Repr	0.68459
197	67X	Heavy Lift Helicopter Repairer	0.68459
198	67Y	AH-1 Attack Helicoptr Repr	0.68459
199	68B	Aircraft Powerplant Repairer	0.68459
200	68D	Aircraft Powertrain Repairer	0.68459
201	68H	Aircraft Pneudraulic Repairer	0.68459
202	62B	Construction Equipment Repr	0.70398
203	36H	Dial/Manual Centrl Office Rep	0.71119
204	63D	SP Field Artilry System Mech	0.84468
205	63E	M1 Abrams Tank System Mech	0.84468
206	63H	Track Vehicle Repairer	0.84468
207	63N	M60A1/A3 Tank System Mechanic	0.84468
208	63T	Bradley System Mechanic	0.84468
209	63Y	Track Vehicle Mechanic	0.84468
210	68G	Aircraft Structural Repairer	0.89747

------------------------------- CLSNUM=6 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
211	61F	Marine Hull Repr	0.50070
212	62G	Quarrying Specialist	0.56695
213	36D	Antenna Installer Specialist	0.70110
214	51R	Interior Electrician	0.79109
215	36E	Cable Splicer	0.87028
216	52G	Transmisson & Distbution Spec	0.90788
217	68M	Aircraft Weapon Systems Repr	1.06768
218	62E	Hvy Construction Equip Op	1.07876
219	62F	Crane Operator	1.07876
220	51K	Plumber	1.08555
221	12E	Atomic Demo Munitions Spec	1.13762
222	45B	Small Arms Repairer	1.13762
223	45D	SP Field Artlry Turret Mech	1.13762
224	45E	M1 ABRAMS Tank Turret Mech	1.13762
225	45K	Tank Turret Repairer	1.13762
226	45L	Artillery Repairer	1.13762
227	45N	M60A1/A3 Tank Turret Mech	1.13762
228	45T	BFVS Turret Mechanic	1.13762
229	55D	Explsve Ordnance Disposl Spec	1.13762
230	55G	Nuclear Weapons Maint Spec	1.13762
231	51B	Carpentry/Masonry Specialist	1.16507

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
SEVEN CLUSTER SOLUTION

------------------------------- CLSNUM=6 --------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
232	52C	Utilities Equipment Repairer	1.19446
233	65D	Railway Car Repairer	1.31183
234	76W	Petroleum Supply Specialist	1.33526
235	51C	Structures Specialist	1.36199
236	00B	Diver	1.38434
237	63J	Quart&Chem Equipment Repairer	1.48484
238	12F	Engineer Tracked Veh Crewman	1.57321
239	51M	Firefighter	1.69732

------------------------------- CLSNUM=7 --------------------------------

OBS	MOS	MOSTITLE	DISTANCE
240	71R	Broadcast Journalist	0.42810
241	82C	Field Artillery Surveyor	0.50668
242	71P	Flight Operations Coordinator	0.69293
243	71Q	Journalist	0.80494
244	71N	Traffic Management Coordntor	0.88675
245	96B	Intelligence Analyst	0.89854
246	98C	EW/SIGINT Analyst	0.89854
247	98J	EW/SIGINT NoncommoIntercept	0.89854
248	96C	Interrogator	0.90186
249	98G	EW/SIGINT Voice Interceptor	0.90186
250	91D	Operating Room Specialist	0.91504
251	93E	Meteorological Observer	0.96454
252	93F	Field Artlry Meteorlogic Spec	0.96454
253	15D	Lance Missile Crewmember	1.00396
254	15E	Pershing Missile Crewmember	1.03815
255	82D	Topographic Surveyor	1.05751
256	71D	Legal Clerk Specialist	1.12012
257	92C	Petroleum Laboratory Spec	1.21438
258	97B	Counterintelligence Agents	1.28063
259	16H	ADA Opertns-Intellignce Assis	1.42687
260	74F	Programmer Analyst	1.43646
261	75C	Personnel Management Spec	1.60056
262	91G	Behavioral Science Specialist	1.92033
263	92B	Medical Laboratory Specialist	2.02246

Table 2.
WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=1

N	Mean	Std Dev	Minimum	Maximum
46	1.1789369	0.4666359	0.3939746	2.4897408

CLSNUM=2

N	Mean	Std Dev	Minimum	Maximum
28	1.3026237	0.5751421	0.4159113	2.4759866

CLSNUM=3

N	Mean	Std Dev	Minimum	Maximum
26	1.3296487	0.4503430	0.6811366	2.3914765

CLSNUM=4

N	Mean	Std Dev	Minimum	Maximum
67	0.5340135	0.2117665	0.3865468	1.0311053

CLSNUM=5

N	Mean	Std Dev	Minimum	Maximum
43	0.6267022	0.1639175	0.1396649	0.8974717

CLSNUM=6

N	Mean	Std Dev	Minimum	Maximum
29	1.1218373	0.2660345	0.5007033	1.6973233

WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

------------------------------ CLSNUM=7 ------------------------------

N	Mean	Std Dev	Minimum	Maximum
24	1.0743455	0.3851911	0.4281011	2.0224638

Table 3.
WARD HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER FACTOR SCORES

OBS	CLSNUM	FREQ	LGEDIST	FACT1	FACT2	FACT3	FACT4
1	1	46	2.48974	-0.53402	-1.01222	-0.71176	-0.16550
2	2	28	2.47599	-0.11661	-1.61539	0.97982	0.17111
3	3	26	2.39148	0.15788	0.24546	-0.45250	1.87808
4	4	67	1.03111	0.88834	0.36909	-0.77713	-0.33684
5	5	43	0.89747	0.29912	0.39088	0.64072	-1.02574
6	6	29	1.69732	0.28593	0.62803	1.81452	0.62703
7	7	24	2.02246	-2.37213	1.06938	-0.45838	0.10496

Appendix C

Average Linkage Hierarchical Cluster Analysis of DOT DATA

- Plot of Cubic Cluster Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
PLOT OF CUBIC CLUSTERING CRITERION BY NUMBER OF CLUSTERS

Plot of _CCC_*_NCL_.   Legend: A = 1 obs, B = 2 obs, etc.

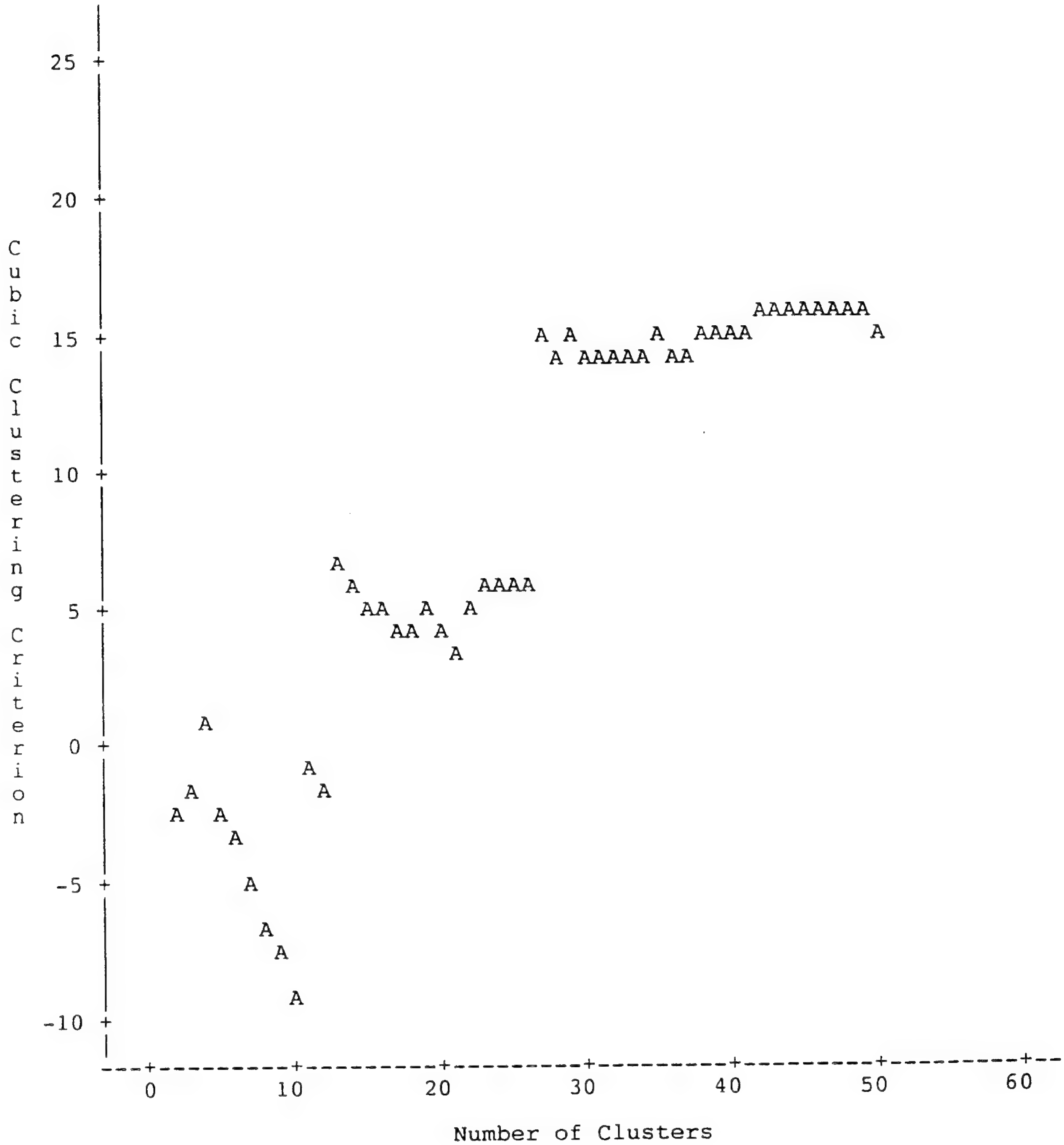


Figure 1. CCC by Number of Clusters

PLOT OF GAMMA VALUES BY NUMBER OF CLUSTERS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

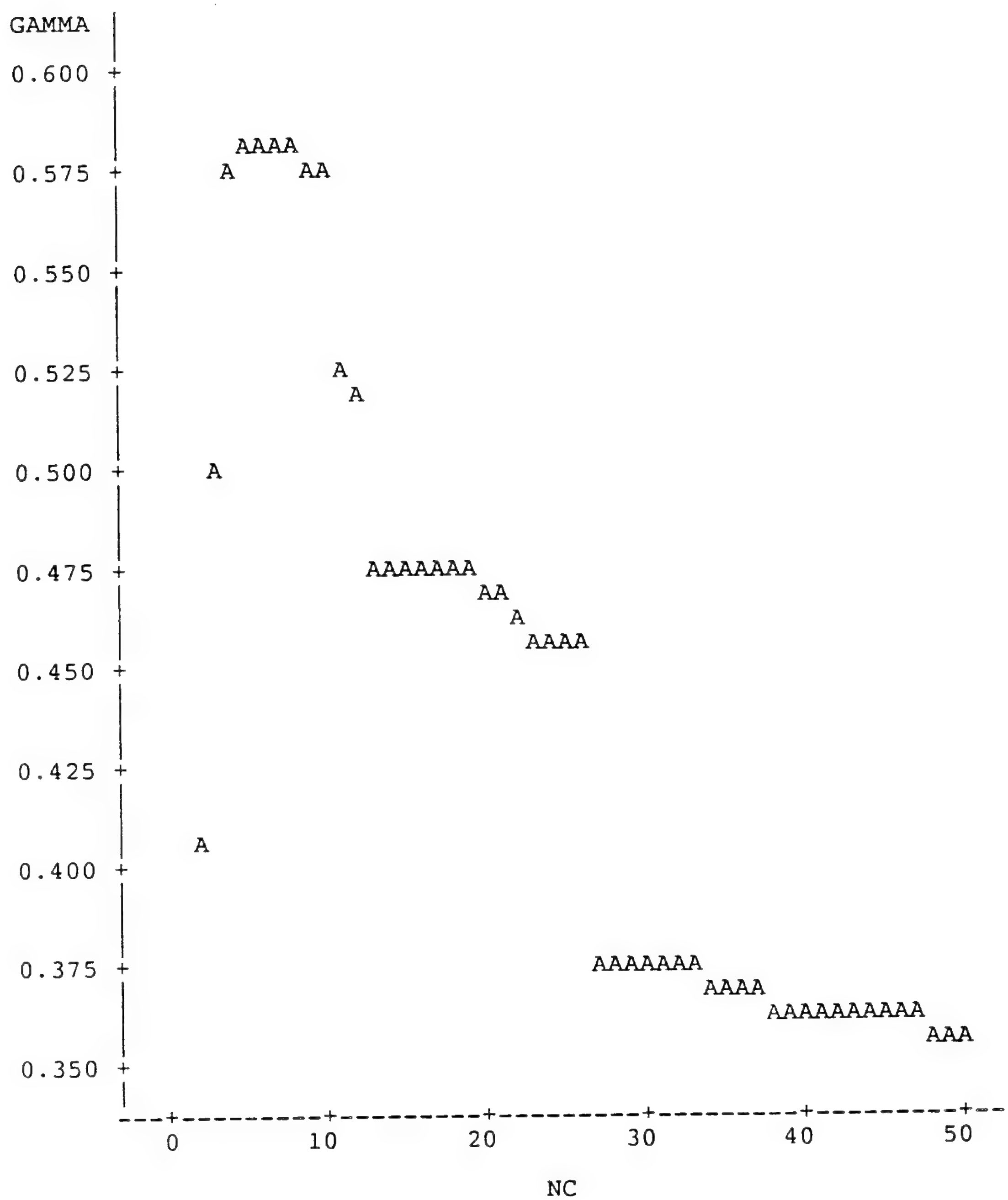


Figure 2. Gamma by Number of Clusters

Table 1.
AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=1 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
1	05C	Radio Teletype Operator	0.68883
2	05D	EW/SIGINT Emitter ID/Locator	0.62500
3	05G	Signal Security Specialist	0.62500
4	05H	EW/SIGINT Intercept-IMC	0.62500
5	05K	EW/SIGINT N-M Interceptor	0.62500
6	16H	ADA Opertns-Intellignce Assis	1.30342
7	65E	Airbrake Repairer	1.43262
8	65H	Locomotive Operator	1.07339
9	71G	Patient Admin Specialist	0.38648
10	71L	Administrative Specialist	0.85473
11	72E	Combat Telecomm Ctr Operator	0.82789
12	72G	Auto Data Telecomm Ctr Oprtor	0.91211
13	74D	Computer/Machine Operator	0.91211
14	75B	Personnel Admin Specialist	0.83621
15	75C	Personnel Management Spec	1.03737
16	75D	Personnel Records Specialist	0.83621
17	75E	Personnel Action Specialist	0.83621
18	75F	Personnel Info Mangmt Spec	0.83621
19	76C	Equip Records & Parts Spec	0.87351
20	76J	Medical Supply Specialist	1.38435
21	76P	Materl Centrl & Acctng Spec	0.76258
22	76V	Mat Storage & Handlng Spec	1.38435
23	76X	Subsistence Supply Specialist	1.38435
24	76Y	Unit Supply Specialist	1.29646
25	91C	Practical Nurse	1.38901
26	91E	Dental Specialist	0.22880
27	91H	Orthopedic Specialist	0.90847
28	91L	Occupational Therapy Spec	0.63446
29	91R	Veterinary Food Inspec Spec	1.20274
30	91S	Environmental Health Spec	1.20274
31	91T	Animal Care Specialist	0.97973
32	91U	Ear, Nose & Throat Specialist	1.26555
33	91Y	Eye Specialist	1.15638
34	93E	Meteorological Observer	1.54980
35	93F	Field Artlry Meteorlogic Spec	1.54980
36	94B	Food Service Specialist	0.86129
37	94F	Hospital Food Service Spec	0.86129

------------------------------ CLSNUM=2 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
38	11B	Infantryman	0.89762
39	11C	Indirect Fire Infantryman	0.72032
40	11H	Heavy Anti-Armor Wpns Inftryman	1.44070
41	11M	Fighting Vehicle Infantryman	1.44070
42	12F	Engineer Tracked Veh Crewman	1.11478
43	13B	Cannon Crewman	0.82993
44	16F	Light Air Def Artillery Crewmbr	1.38202
45	16L	Sgt York Air Def Gun Crwmbr	1.38202
46	16R	ADA Short Range Gunry Crwmbr	1.38202
47	19E	M48-M60 Armor Crewman	0.64748

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

--- CLSNUM=2 ---
(continued)

OBS	MOS	MOSTITLE	DISTANCE
48	19K	M1 ABRAMS Armor Crewman	0.64748
49	22N	NIKEHER MissileLauncherRepair	1.04464
50	24L	ImpHAWK Launch&Mech Sys Repr	1.04464
51	27H	HAWK Firing Section Reparirer	1.04464
52	46N	PERSHING ElecMechcal Repairer	1.04464
53	54C	Smoke Operation Specialist	0.33643
54	61B	Watercraft Operator	0.88630
55	62H	Concrete & Asphalt Equip Op	0.91755
56	62J	General Construc Equip Op	0.73337
57	64C	Motor Transport Operator	0.87028
58	65J	Train Crewmemeber	1.31280

--- CLSNUM=3 ---

OBS	MOS	MOSTITLE	DISTANCE
59	13E	Cannon Fire Direction Speclst	0.50207
60	13F	Fire Support Specialist	1.48674
61	15J	MLRS/LANCE Ops/FireDir Spec	0.50207
62	17L	Aerial Sensor Specialist	1.24117
63	42C	Orthotic Specialist	0.73201
64	71C	Exec Administrative Assistant	0.95289
65	71E	Court Reporter	1.44560
66	81C	Cartographer	1.34276
67	82B	Construction Surveyor	1.50225
68	91F	Psychiatric Specialist	1.48147
69	91J	Physical Therapy Specialist	1.19549
70	91N	Cardiac Specialist	0.68008
71	91P	X-Ray Specialist	0.82596
72	93H	Air Traffic Control Tower Op	0.72894
73	93J	Air Traff Cntrl Radar Contlr	0.72894
74	95B	Military Police	1.23638
75	95C	Correctional Specialist	1.55959
76	96D	Imagery Analyst	1.05799
77	96H	Aerial Intell Spec	1.24117

--- CLSNUM=4 ---

OBS	MOS	MOSTITLE	DISTANCE
78	05B	Radio Operator	0.72867
79	21L	PERSHING Electronics Repairer	0.71564
80	22L	NIKE Test Equipment Repairer	0.71564
81	23N	NIKE Track Radar Repairer	0.71564
82	23U	NIKE Radar-Simulator Repairer	0.71564
83	24C	Improvd HAWK Firng Sect Mech	0.71564
84	24E	Improvd HAWK Fire Contrl Mech	0.71564
85	24G	Imp HAWK Inform CoorCentMech	0.71564
86	24H	Improvd HAWK Fire Contrl Repr	0.71564
87	24J	Improved HAWK Pulse Radar Rep	0.71564

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=4 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
88	24K	ImpHAWK Cont-Wave Radar Repr	0.71564
89	24P	Defense Acq Radar Mechanic	0.71564
90	24Q	NIKE-HERCULES Fire Contrl Mec	0.71564
91	24U	HERCULES Electronic Mechanic	0.71564
92	24W	Sgt York Air Def Gun Syst Mec	0.71564
93	25J		0.71564
94	26B	Weapons Support Radar Repr	0.71564
95	26D	Ground Cntrl Approch Rdar Rep	0.71564
96	26E	Aerial Surv Sensor Repairer	0.71564
97	26F	Aerial Photo-Activ Sensor Rep	0.42556
98	26H	Air Defense Radar Repairer	0.71564
99	26L	Tactical Microwave Syst Repr	1.29866
100	26M	Aerial Surveillance Radar Repr	0.71564
101	26N	Aerial Surveillance Infrared Repr	0.41097
102	26Q	Tact Satell/Microwave Syst Op	0.72867
103	26R	Strategic Microwave Syst Op	0.72867
104	26T	Radio/TV Systems Specialist	0.69233
105	26V	Strategic Microwave Syst Repr	1.29866
106	26Y	SATCOM Equipment Repairer	0.71564
107	27B	LandCombat SystemTestSpecial	0.71564
108	27E	TOW/DRAGON Repairer	0.71564
109	27F	VULCAN Repairer	0.71564
110	27G	CHAPARRAL/REDEYE Repairer	0.71564
111	27L	LANCE System Repairer	0.71564
112	27M	MLRS Repairer	0.71564
113	27N	Forwrd Area Alerting Rdar Rep	0.71564
114	27P	SgtYork Radar/Electron Repr	0.71564
115	27Q	SgtYork Test Specialist	0.71564
116	31E	Field Radio Repairer	0.41097
117	31M	Multichannel Commo Equip Op	0.72867
118	31S	Field General Comsec Repairer	0.41097
119	31T	Field Systems Comsec Repairer	0.41097
120	31V	Tactical Commo Syst Op/Mech	0.72867
121	32D	Station Technical Controller	0.41097
122	32F	Fixed Ciphony Repairer	0.71564
123	32G	Fixed Crypto Equip Repairer	0.71564
124	32H	Fixed Station Radio Repairer	1.29866
125	33S	EW/Intercept Sys Repr	0.71564
126	34B	Punch Card Machine Operator	0.36624
127	34C	Decen Auto Serv Supp Systm	0.71564
128	34E	NCR 500 Computer Repairer	0.71564
129	34F	Digt Subsc Message Switch Equip	0.71564
130	34H	Auto Digt Message Switch Equip	0.71564
131	34Y	Field Artlry TactFire Repair	0.71564
132	35B	Electronics Instrument Repr	0.71564
133	35C	Automatic Test Equip Repairer	0.71564
134	35E	Special Elec Devices Repairer	0.59969
135	35F	Nuclear Weapons Electronics Specialist	0.71564
136	35G	Biomedical Equipment Spec	0.55607
137	35K	Avionic Mechanic	0.96493
138	35L	Avionic Commo Equip Repairer	0.96493
139	35M	Avionic Nav/FlightContEq Repr	0.96493

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=4 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
140	35R	Avionic Special Equip Repr	0.96493
141	41C	Fire Contrl Instru Rep Spec	0.59969
142	41E	Audio/Visual Equip Repairer	0.36362
143	41G	Aerial Surveillance Photo Equip Repr	0.42556
144	41J	Office Machine Repairer	0.34791
145	42D	Dental Laboratory Specialist	0.54935
146	42E	Optical Laboratory Spec	0.71672
147	44B	Metalworker	1.24081
148	44E	Machinist	1.04048
149	45G	Fire Control System Repairer	0.71564
150	52D	Power Generator Equip Repr	1.23937
151	61C	Watercraft Engineer	1.27941
152	63D	SP Field Artilry System Mech	1.22580
153	63E	M1 Abrams Tank System Mech	1.22580
154	63H	Track Vehicle Repairer	1.22580
155	63N	M60A1/A3 Tank System Mechanic	1.22580
156	63T	Bradley System Mechanic	1.22580
157	63Y	Track Vehicle Mechanic	1.22580
158	65B	Locomotive Repairer	1.30153
159	68F	Aircraft Electrician	1.02167
160	81B	Technical Drafting Specialist	0.57667
161	83E	Photo & Layout Specialist	1.27465
162	83F	Photolithographer	1.23886
163	91Q	Pharmacy Specialist	0.77180
164	91V	Respiratory Specialist	0.85552
165	92D	Chemical Laboratory Spec	0.77180

------------------------------ CLSNUM=5 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
166	31J	Teletypewriter Repairer	0.70663
167	36C	Wire System Instll/Operator	0.91918
168	36D	Antenna Installer Specialist	0.94810
169	36E	Cable Splicer	1.03666
170	36H	Dial/Manual Centrl Office Rep	0.95622
171	36L	Trans ElectSwitchSys Rep	0.70663
172	41B	Topographic Instr Rep Spec	0.80308
173	51B	Carpentry/Masonry Specialist	0.53007
174	51C	Structures Specialist	0.69036
175	51K	Plumber	0.63851
176	51R	Interior Electrician	1.23163
177	52C	Utilities Equipment Repairer	0.83445
178	62B	Construction Equipment Repr	0.44280
179	62E	Hvy Construction Equip Op	0.70250
180	62F	Crane Operator	0.70250
181	63B	Light Wheel Vehicle Mechanic	0.38038
182	63G	Fuel & Elec System Repairer	0.38038
183	63J	Quart&Chem Equipment Repairer	0.91764
184	63S	Heavy Wheel Vehicle Mechanic	0.38038
185	63W	Wheel Vehicle Repairer	0.38038

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=5 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
186	65D	Railway Car Repairer	0.67288
187	67G	Utility Airplane Repairer	0.40190
188	67H	Observation Airplane Repairer	0.40190
189	67N	Utility Helicopter Repairer	0.40190
190	67T	Tact Transp Helicoptr Repr	0.40190
191	67U	Medium Helicopter Repairer	0.40190
192	67V	Observ/Scout Helicoptr Repr	0.40190
193	67X	Heavy Lift Helicopter Repairer	0.40190
194	67Y	AH-1 Attack Helicoptr Repr	0.40190
195	68B	Aircraft Powerplant Repairer	0.40190
196	68D	Aircraft Powertrain Repairer	0.40190
197	68G	Aircraft Structural Repairer	0.88402
198	68H	Aircraft Pneudraulic Repairer	0.40190
199	76W	Petroleum Supply Specialist	0.79126

------------------------------ CLSNUM=6 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
200	31N	Tactical Circuit Controller	1.02945
201	35H	Calibration Specialist	1.20839
202	36K	Tactical Wire Operations Specialist	0.94041
203	36M	Wire Systems Operator	1.02945
204	43E	Parachute Rigger	0.96198
205	43M	Fabric Repair Specialist	0.70664
206	53B	Industrial Gas Prod Specialist	0.79970
207	57F	Graves Registration Spec	1.55235
208	57H	Cargo Specialist	1.62742
209	73C	Finance Specialist	0.79399
210	73D	Accounting Specialist	0.62781
211	74B	Card and Tape Writer	0.84376

------------------------------ CLSNUM=7 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
212	00B	Diver	1.84317
213	12E	Atomic Demo Munitions Spec	0.45310
214	45B	Small Arms Repairer	0.45310
215	45D	SP Field Artlry Turret Mech	0.45310
216	45E	M1 ABRAMS Tank Turret Mech	0.45310
217	45K	Tank Turret Repairer	0.45310
218	45L	Artillery Repairer	0.45310
219	45N	M60A1/A3 Tank Turret Mech	0.45310
220	45T	BFVS Turret Mechanic	0.45310
221	51M	Firefighter	1.85564
222	52G	Transmisson & Distbution Spec	0.81101
223	55D	Explsve Ordnance Disposl Spec	0.45310
224	55G	Nuclear Weapons Maint Spec	0.45310
225	61F	Marine Hull Repr	1.02417

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

---------------------------------- CLSNUM=7 ----------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
226	62G	Quarrying Specialist	0.74560
227	68M	Aircraft Weapon Systems Repr	0.84169

---------------------------------- CLSNUM=8 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
228	12B	Combat Engineer	1.66756
229	12C	Bridge Crewman	1.07971
230	51N	Water Treatment Specialist	1.02144
231	54E	NBC Specialist	1.10058
232	55B	Ammunition Specialist	0.76453
233	55R	Ammo Stock Control&Acct Spec	0.76453
234	57E	Laundry & Bath Specialist	0.97729

---------------------------------- CLSNUM=9 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
235	81E	Illustrator	0.93167
236	84B	Still Photographic Specialist	0.31056
237	84C	Motion Picture Specialist	0.31056
238	84F	Audio/TV Specialist	0.31056

---------------------------------- CLSNUM=10 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
239	15D	Lance Missile Crewmember	0.97988
240	15E	Pershing Missile Crewmember	1.02455
241	71D	Legal Clerk Specialist	1.11196
242	71N	Traffic Management Coordntor	0.91491
243	71P	Flight Operations Coordinator	0.66499
244	71Q	Journalist	0.77575
245	71R	Broadcast Journalist	0.44743
246	74F	Programmer Analyst	1.39501
247	82C	Field Artillery Surveyor	0.54173
248	82D	Topographic Surveyor	1.09013
249	91D	Operating Room Specialist	0.91205
250	92C	Petroleum Laboratory Spec	1.19312
251	96B	Intelligence Analyst	0.87434
252	96C	Interrogator	0.94608
253	97B	Counterintelligence Agents	1.30024
254	98C	EW/SIGINT Analyst	0.87434
255	98G	EW/SIGINT Voice Interceptor	0.94608
256	98J	EW/SIGINT NoncommoIntercept	0.87434

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
THIRTEEN CLUSTER SOLUTION

----------------------------------- CLSNUM=11 -----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
257	03C	Physical Activities Spec	1.52619
258	17M	Remote Sensor Specialist	0.88760
259	19D	Cavalry Scout	0.94540
260	91A	Medical Specialist	0.76536

----------------------------------- CLSNUM=12 -----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
261	91G	Behavioral Science Specialist	0.60523
262	92B	Medical Laboratory Specialist	0.60523

----------------------------------- CLSNUM=13 -----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
263	17C	Field Artlry Target Acq Spec	0

Table 2.
AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=1

N	Mean	Std Dev	Minimum	Maximum
37	0.9770125	0.3243709	0.2288037	1.5498006

CLSNUM=2

N	Mean	Std Dev	Minimum	Maximum
21	1.0057306	0.3048496	0.3364318	1.4406966

CLSNUM=3

N	Mean	Std Dev	Minimum	Maximum
19	1.0759780	0.3577521	0.5020740	1.5595874

CLSNUM=4

N	Mean	Std Dev	Minimum	Maximum
88	0.7854625	0.2522928	0.3479092	1.3015272

CLSNUM=5

N	Mean	Std Dev	Minimum	Maximum
34	0.6193410	0.2433421	0.3803835	1.2316280

CLSNUM=6

N	Mean	Std Dev	Minimum	Maximum
12	1.0101132	0.3132278	0.6278085	1.6274205

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=7

N	Mean	Std Dev	Minimum	Maximum
16	0.7282702	0.4746015	0.4531041	1.8556400

CLSNUM=8

N	Mean	Std Dev	Minimum	Maximum
7	1.0536626	0.3038122	0.7645306	1.6675560

CLSNUM=9

N	Mean	Std Dev	Minimum	Maximum
4	0.4658368	0.3105579	0.3105579	0.9316736

CLSNUM=10

N	Mean	Std Dev	Minimum	Maximum
18	0.9370519	0.2400782	0.4474304	1.3950147

CLSNUM=11

N	Mean	Std Dev	Minimum	Maximum
4	1.0311399	0.3384609	0.7653635	1.5261918

CLSNUM=12

N	Mean	Std Dev	Minimum	Maximum
2	0.6052264	0	0.6052264	0.6052264

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

------------------------------- CLSNUM=13 -------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3.
AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER FACTOR SCORES

OBS	CLSNUM	FREQ	LGEDIST	FACT1	FACT2	FACT3	FACT4
1	1	37	1.54980	-0.73462	-0.46541	-0.64335	-0.20651
2	2	21	1.44070	0.15738	-1.57081	1.01514	0.55295
3	3	19	1.55959	-0.01389	0.01795	-0.55174	1.40342
4	4	88	1.30153	0.74539	0.33458	-0.53981	-0.55015
5	5	34	1.23163	0.24056	0.52238	1.27509	-0.64453
6	6	12	1.62742	-0.24017	-2.26975	-0.83617	-0.06792
7	7	16	1.85564	0.41263	0.84044	1.90488	1.30256
8	8	7	1.66756	-0.87829	-1.47271	1.13371	-1.12757
9	9	4	0.93167	0.76575	1.20575	-1.01800	2.73625
10	10	18	1.39501	-2.36278	1.13633	-0.48111	0.09856
11	11	4	1.52619	0.07425	0.06550	0.45950	3.07850
12	12	2	0.60523	-3.94700	2.03100	-0.14600	0.24700
13	13	1	0.00000	-0.12400	-2.83600	-0.04300	2.06700

Appendix D

K-Means Cluster Analysis of DOT Data

- Plot of Cubic Cluster Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

K-MEANS CLUSTER ANALYSIS OF DOT DATA
PLOT OF CUBIC CLUSTERING CRITERION BY NUMBER OF CLUSTERS

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.

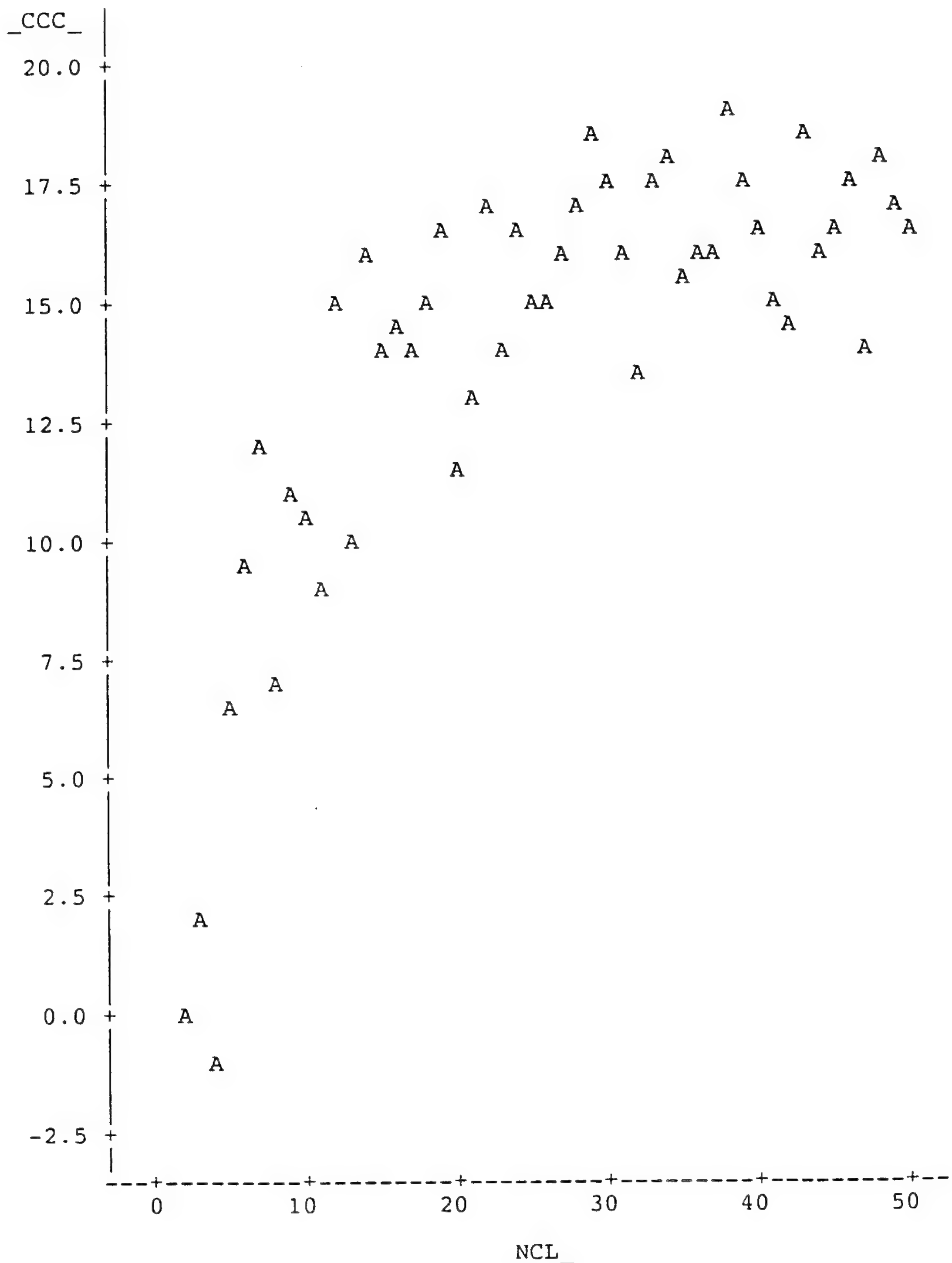


Figure 1. CCC by Number of Clusters

PLOT OF GAMMA VALUES BY NUMBER OF CLUSTERS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

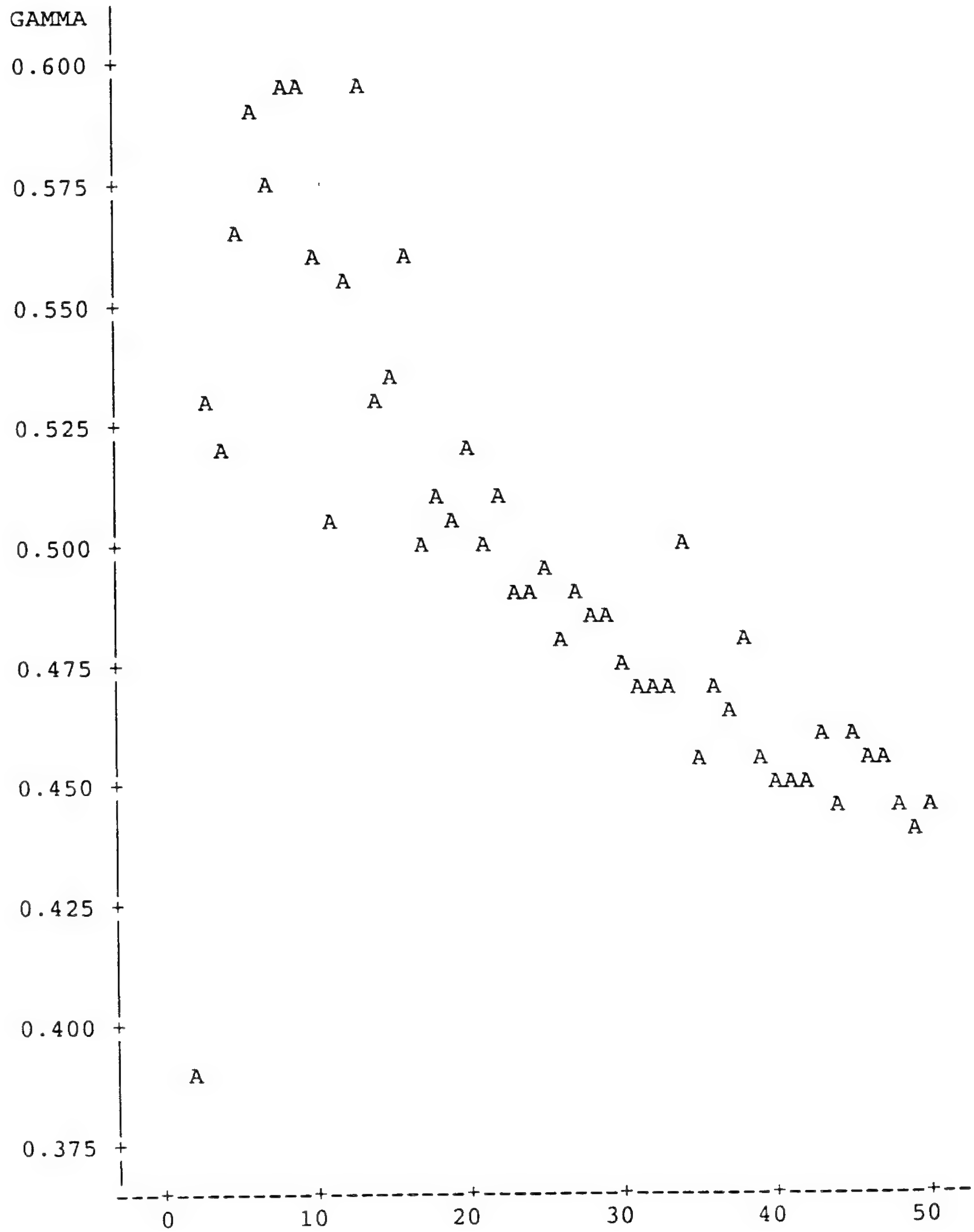


Figure 2. Gamma by Number of Clusters

Table 1.
K-MEANS CLUSTER ANALYSIS OF DOT DATA
SIX CLUSTER SOLUTION

Cluster=1

OBS	MOS	MOSTITLE	DISTANCE
1	13B	Cannon Crewman	0.56325
2	19E	M48-M60 Armor Crewman	0.56489
3	19K	M1 ABRAMS Armor Crewman	0.56489
4	11C	Indirect Fire Infantryman	0.61653
5	54C	Smoke Operation Specialist	0.62656
6	62H	Concrete & Asphalt Equip Op	0.75090
7	62J	General Construc Equip Op	0.82244
8	11B	Infantryman	0.85175
9	64C	Motor Transport Operator	0.85847
10	61B	Watercraft Operator	1.01768
11	22N	NIKEHER MissileLauncherRepair	1.24406
12	24L	ImpHAWK Launch&Mech Sys Repr	1.24406
13	27H	HAWK Firing Section Reparirer	1.24406
14	46N	PERSHING ElecMechcal Repairer	1.24406
15	12F	Engineer Tracked Veh Crewman	1.30749
16	65J	Train Crewmemeber	1.30784
17	16F	Light Air Def Artillery Crewmbr	1.32131
18	16L	Sgt York Air Def Gun Crwmbr	1.32131
19	16R	ADA Short Range Gunry Crwmbr	1.32131
20	12B	Combat Engineer	1.39214
21	11H	Heavy Anti-Armor Wpns Inftryman	1.56895
22	11M	Fighting Vehicle Infantryman	1.56895
23	57F	Graves Registration Spec	1.75939
24	57H	Cargo Specialist	2.01581
25	12C	Bridge Crewman	2.04076
26	57E	Laundry & Bath Specialist	2.14249
27	17C	Field Artlry Target Acq Spec	2.28251

Cluster=2

OBS	MOS	MOSTITLE	DISTANCE
28	75B	Personnel Admin Specialist	0.27386
29	75D	Personnel Records Specialist	0.27386
30	75E	Personnel Action Specialist	0.27386
31	75F	Personnel Info Mangmt Spec	0.27386
32	72E	Combat Telecomm Ctr Operator	0.48988
33	91H	Orthopedic Specialist	0.57232
34	71G	Patient Admin Specialist	0.58428
35	91U	Ear, Nose & Throat Specialist	0.68814
36	71L	Administrative Specialist	0.75894
37	72G	Auto Data Telecomm Ctr Oprtor	0.79380
38	74D	Computer/Machine Operator	0.79380
39	05C	Radio Teletype Operator	0.85682
40	91L	Occupational Therapy Spec	0.86935
41	76C	Equip Records & Parts Spec	0.86987
42	91E	Dental Specialist	0.89611
43	73D	Accounting Specialist	0.99140
44	05D	EW/SIGINT Emitter ID/Locator	1.07801
45	05G	Signal Security Specialist	1.07801
46	05H	EW/SIGINT Intercept-IMC	1.07801
47	05K	EW/SIGINT N-M Interceptor	1.07801

K-MEANS CLUSTER ANALYSIS OF DOT DATA
SIX CLUSTER SOLUTION

------------------------------ Cluster=2 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
48	91F	Psychiatric Specialist	1.08129
49	74B	Card and Tape Writer	1.10758
50	91Y	Eye Specialist	1.12587
51	91T	Animal Care Specialist	1.12657
52	91C	Practical Nurse	1.15842
53	76P	Materl Centrl & Acctng Spec	1.18202
54	73C	Finance Specialist	1.30623
55	31N	Tactical Circuit Controller	1.31265
56	36M	Wire Systems Operator	1.31265
57	35H	Calibration Specialist	1.32295
58	36K	Tactical Wire Operations Specialist	1.33765
59	75C	Personnel Management Spec	1.38518
60	65E	Airbrake Repairer	1.40611
61	43E	Parachute Rigger	1.61934
62	76Y	Unit Supply Specialist	1.62667
63	43M	Fabric Repair Specialist	1.69049
64	91J	Physical Therapy Specialist	1.69394
65	95C	Correctional Specialist	1.94392
66	53B	Industrial Gas Prod Specialist	2.04927
67	71E	Court Reporter	2.08242

------------------------------ Cluster=3 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
68	21L	PERSHING Electronics Repairer	0.38213
69	22L	NIKE Test Equipment Repairer	0.38213
70	23N	NIKE Track Radar Repairer	0.38213
71	23U	NIKE Radar-Simulator Repairer	0.38213
72	24C	Improvd HAWK Firng Sect Mech	0.38213
73	24E	Improvd HAWK Fire Contrl Mech	0.38213
74	24G	Imp HAWK Inform CoorCentMech	0.38213
75	24H	Improvd HAWK Fire Contrl Repr	0.38213
76	24J	Improved HAWK Pulse Radar Rep	0.38213
77	24K	ImpHAWK Cont-Wave Radar Repr	0.38213
78	24P	Defense Acq Radar Mechanic	0.38213
79	24Q	NIKE-HERCULES Fire Contrl Mec	0.38213
80	24U	HERCULES Electronic Mechanic	0.38213
81	24W	Sgt York Air Def Gun Syst Mec	0.38213
82	25J		0.38213
83	26B	Weapons Support Radar Repr	0.38213
84	26D	Ground Cntrl Approch Rdar Rep	0.38213
85	26E	Aerial Surv Sensor Repairer	0.38213
86	26H	Air Defense Radar Repairer	0.38213
87	26M	Aerial Surveillance Radar Repr	0.38213
88	26Y	SATCOM Equipment Repairer	0.38213
89	27B	LandCombat SystemTestSpecial	0.38213
90	27E	TOW/DRAGON Repairer	0.38213
91	27F	VULCAN Repairer	0.38213
92	27G	CHAPARRAL/REDEYE Repairer	0.38213
93	27L	LANCE System Repairer	0.38213

K-MEANS CLUSTER ANALYSIS OF DOT DATA
SIX CLUSTER SOLUTION

------------------------------ Cluster=3 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
94	27M	MLRS Repairer	0.38213
95	27N	Forwrd Area Alerting Rdar Rep	0.38213
96	27P	SgtYork Radar/Electron Repr	0.38213
97	27Q	SgtYork Test Specialist	0.38213
98	32F	Fixed Ciphony Repairer	0.38213
99	32G	Fixed Crypto Equip Repairer	0.38213
100	33S	EW/Intercept Sys Repr	0.38213
101	34C	Decen Auto Serv Supp Systm	0.38213
102	34E	NCR 500 Computer Repairer	0.38213
103	34F	Digt Subsc Message Switch Equip	0.38213
104	34H	Auto Digt Message Switch Equip	0.38213
105	34Y	Field Artlry TactFire Repair	0.38213
106	35B	Electronics Instrument Repr	0.38213
107	35C	Automatic Test Equip Repairer	0.38213
108	35F	Nuclear Weapons Electronics Specialist	0.38213
109	45G	Fire Control System Repairer	0.38213
110	81B	Technical Drafting Specialist	0.48847
111	41E	Audio/Visual Equip Repairer	0.53671
112	42D	Dental Laboratory Specialist	0.60067
113	41J	Office Machine Repairer	0.64934
114	34B	Punch Card Machine Operator	0.66112
115	35G	Biomedical Equipment Spec	0.67105
116	26F	Aerial Photo-Activ Sensor Rep	0.72574
117	41G	Aerial Surveillance Photo Equip Repr	0.72574
118	26N	Aerial Surveillance Infrared Repr	0.78816
119	31E	Field Radio Repairer	0.78816
120	31S	Field General Comsec Repairer	0.78816
121	31T	Field Systems Comsec Repairer	0.78816
122	32D	Station Technical Controller	0.78816
123	35E	Special Elec Devices Repairer	0.79999
124	41C	Fire Contrl Instru Rep Spec	0.79999
125	91Q	Pharmacy Specialist	0.81372
126	05B	Radio Operator	0.87051
127	26Q	Tact Satell/Microwave Syst Op	0.87051
128	26R	Strategic Microwave Syst Op	0.87051
129	31M	Multichannel Commo Equip Op	0.87051
130	31V	Tactical Commo Syst Op/Mech	0.87051
131	92D	Chemical Laboratory Spec	0.97151
132	91V	Respiratory Specialist	1.01279
133	26T	Radio/TV Systems Specialist	1.01825
134	81C	Cartographer	1.05268
135	42E	Optical Laboratory Spec	1.08521
136	91P	X-Ray Specialist	1.33936
137	96D	Imagery Analyst	1.34867
138	42C	Orthotic Specialist	1.48671

------------------------------ Cluster=4 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
139	65B	Locomotive Repairer	0.37404

K-MEANS CLUSTER ANALYSIS OF DOT DATA
SIX CLUSTER SOLUTION

------------------------------- Cluster=4 -------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
140	83F	Photolithographer	0.41320
141	61C	Watercraft Engineer	0.44315
142	31J	Teletypewriter Repairer	0.46117
143	36L	Trans ElectSwitchSys Rep	0.46117
144	62B	Construction Equipment Repr	0.47033
145	41B	Topographic Instr Rep Spec	0.47371
146	63B	Light Wheel Vehicle Mechanic	0.50977
147	63G	Fuel & Elec System Repairer	0.50977
148	63S	Heavy Wheel Vehicle Mechanic	0.50977
149	63W	Wheel Vehicle Repairer	0.50977
150	36C	Wire System Instll/Operator	0.53297
151	44B	Metalworker	0.53415
152	52D	Power Generator Equip Repr	0.64708
153	83E	Photo & Layout Specialist	0.65179
154	67G	Utility Airplane Repairer	0.70452
155	67H	Observation Airplane Repairer	0.70452
156	67N	Utility Helicopter Repairer	0.70452
157	67T	Tact Transp Helicoptr Repr	0.70452
158	67U	Medium Helicopter Repairer	0.70452
159	67V	Observ/Scout Helicoptr Repr	0.70452
160	67X	Heavy Lift Helicopter Repairer	0.70452
161	67Y	AH-1 Attack Helicoptr Repr	0.70452
162	68B	Aircraft Powerplant Repairer	0.70452
163	68D	Aircraft Powertrain Repairer	0.70452
164	68H	Aircraft Pneudraulic Repairer	0.70452
165	36H	Dial/Manual Centrl Office Rep	0.77303
166	68F	Aircraft Electrician	0.78236
167	68G	Aircraft Structural Repairer	0.78626
168	26L	Tactical Microwave Syst Repr	0.80416
169	26V	Strategic Microwave Syst Repr	0.80416
170	32H	Fixed Station Radio Repairer	0.80416
171	44E	Machinist	0.85020
172	35K	Avionic Mechanic	0.86800
173	35L	Avionic Commo Equip Repairer	0.86800
174	35M	Avionic Nav/FlightContEq Repr	0.86800
175	35R	Avionic Special Equip Repr	0.86800
176	63D	SP Field Artilry System Mech	0.98201
177	63E	M1 Abrams Tank System Mech	0.98201
178	63H	Track Vehicle Repairer	0.98201
179	63N	M60A1/A3 Tank System Mechanic	0.98201
180	63T	Bradley System Mechanic	0.98201
181	63Y	Track Vehicle Mechanic	0.98201
182	62E	Hvy Construction Equip Op	1.02897
183	62F	Crane Operator	1.02897
184	51C	Structures Specialist	1.02999
185	65D	Railway Car Repairer	1.07383
186	51B	Carpentry/Masonry Specialist	1.08911
187	94B	Food Service Specialist	1.10301
188	94F	Hospital Food Service Spec	1.10301
189	51K	Plumber	1.11281
190	76W	Petroleum Supply Specialist	1.17196
191	65H	Locomotive Operator	1.21611

K-MEANS CLUSTER ANALYSIS OF DOT DATA
SIX CLUSTER SOLUTION

------------------------------ Cluster=4 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
192	51N	Water Treatment Specialist	1.24944
193	36E	Cable Splicer	1.30013
194	52C	Utilities Equipment Repairer	1.37121
195	36D	Antenna Installer Specialist	1.40211
196	63J	Quart&Chem Equipment Repairer	1.46175
197	55B	Ammunition Specialist	1.65878
198	55R	Ammo Stock Control&Acct Spec	1.65878
199	91R	Veterinary Food Inspec Spec	1.66536
200	91S	Environmental Health Spec	1.66536
201	51R	Interior Electrician	1.69019
202	76J	Medical Supply Specialist	1.74966
203	76V	Mat Storage & Handlng Spec	1.74966
204	76X	Subsistence Supply Specialist	1.74966
205	61F	Marine Hull Repr	1.82515
206	54E	NBC Specialist	2.01378
207	00B	Diver	2.22370
208	51M	Firefighter	2.95684

------------------------------ Cluster=5 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
209	68M	Aircraft Weapon Systems Repr	0.99098
210	13E	Cannon Fire Direction Speclst	1.03622
211	15J	MLRS/LANCE Ops/FireDir Spec	1.03622
212	95B	Military Police	1.24848
213	12E	Atomic Demo Munitions Spec	1.39688
214	45B	Small Arms Repairer	1.39688
215	45D	SP Field Artlry Turret Mech	1.39688
216	45E	M1 ABRAMS Tank Turret Mech	1.39688
217	45K	Tank Turret Repairer	1.39688
218	45L	Artillery Repairer	1.39688
219	45N	M60A1/A3 Tank Turret Mech	1.39688
220	45T	BFVS Turret Mechanic	1.39688
221	55D	Explsve Ordnance Disposl Spec	1.39688
222	55G	Nuclear Weapons Maint Spec	1.39688
223	82B	Construction Surveyor	1.41130
224	13F	Fire Support Specialist	1.45644
225	52G	Transmisson & Distbution Spec	1.49405
226	17L	Aerial Sensor Specialist	1.53110
227	96H	Aerial Intell Spec	1.53110
228	17M	Remote Sensor Specialist	1.53143
229	93H	Air Traffic Control Tower Op	1.59766
230	93J	Air Traff Cntrl Radar Contlr	1.59766
231	84B	Still Photographic Specialist	1.62190
232	84C	Motion Picture Specialist	1.62190
233	84F	Audio/TV Specialist	1.62190
234	91A	Medical Specialist	1.66088
235	19D	Cavalry Scout	1.68463
236	91N	Cardiac Specialist	1.68962
237	62G	Quarrying Specialist	1.83390

K-MEANS CLUSTER ANALYSIS OF DOT DATA
SIX CLUSTER SOLUTION

------------------------------ Cluster=5 ------------------------------
(continued)

OBS	MOS	MOSTITLE	DISTANCE
238	03C	Physical Activities Spec	1.95916
239	71C	Exec Administrative Assistant	1.96480
240	81E	Illustrator	2.74549

------------------------------ Cluster=6 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
241	71R	Broadcast Journalist	0.44777
242	82C	Field Artillery Surveyor	0.54370
243	71P	Flight Operations Coordinator	0.72971
244	71Q	Journalist	0.74334
245	91D	Operating Room Specialist	0.90323
246	96B	Intelligence Analyst	0.90354
247	98C	EW/SIGINT Analyst	0.90354
248	98J	EW/SIGINT NoncommoIntercept	0.90354
249	15D	Lance Missile Crewmember	0.93467
250	71N	Traffic Management Coordntor	0.93930
251	96C	Interrogator	0.95773
252	98G	EW/SIGINT Voice Interceptor	0.95773
253	93E	Meteorological Observer	1.00534
254	93F	Field Artlry Meteorlogic Spec	1.00534
255	82D	Topographic Surveyor	1.03065
256	15E	Pershing Missile Crewmember	1.03820
257	71D	Legal Clerk Specialist	1.10301
258	92C	Petroleum Laboratory Spec	1.15990
259	97B	Counterintelligence Agents	1.30007
260	74F	Programmer Analyst	1.43659
261	16H	ADA Opertns-Intellignce Assis	1.49071
262	91G	Behavioral Science Specialist	1.86335
263	92B	Medical Laboratory Specialist	1.95765

Table 2.
K-MEANS CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

Cluster=1

N	Mean	Std Dev	Minimum	Maximum
27	1.2431065	0.5029299	0.5632547	2.2825120

Cluster=2

N	Mean	Std Dev	Minimum	Maximum
40	1.0860850	0.4692085	0.2738581	2.0824204

Cluster=3

N	Mean	Std Dev	Minimum	Maximum
71	0.5793054	0.2805117	0.3821345	1.4867081

Cluster=4

N	Mean	Std Dev	Minimum	Maximum
70	1.0081314	0.4945853	0.3740368	2.9568440

Cluster=5

N	Mean	Std Dev	Minimum	Maximum
32	1.5261133	0.3161746	0.9909758	2.7454893

Cluster=6

N	Mean	Std Dev	Minimum	Maximum
23	1.0547211	0.3602128	0.4477715	1.9576463

Table 3.
K-MEANS CLUSTER ANALYSIS OF DOT DATA
MEAN CLUSTER FACTOR SCORES

OBS	FREQ	LGEDIST	NEAREST	FACT1	FACT2	FACT3	FACT4
1	27	2.28251	2	-0.02978	-1.81959	0.98244	0.38204
2	40	2.08242	3	-0.52085	-1.03655	-0.86955	0.13608
3	71	1.48671	4	0.86872	0.36625	-0.77555	-0.26866
4	70	2.95684	3	0.07260	0.28111	0.82223	-0.84264
5	32	2.74549	3	0.32709	0.60250	0.49000	1.87741
6	23	1.95765	2	-2.41622	1.11448	-0.42974	0.09830

Appendix E

Ward Hierarchical Cluster Analysis of Project A Data

- Plot of Gamma Values by
  Number of Clusters

- Composition of Clusters

PLOT OF GAMMA VALUES BY NUMBER OF CLUSTERS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

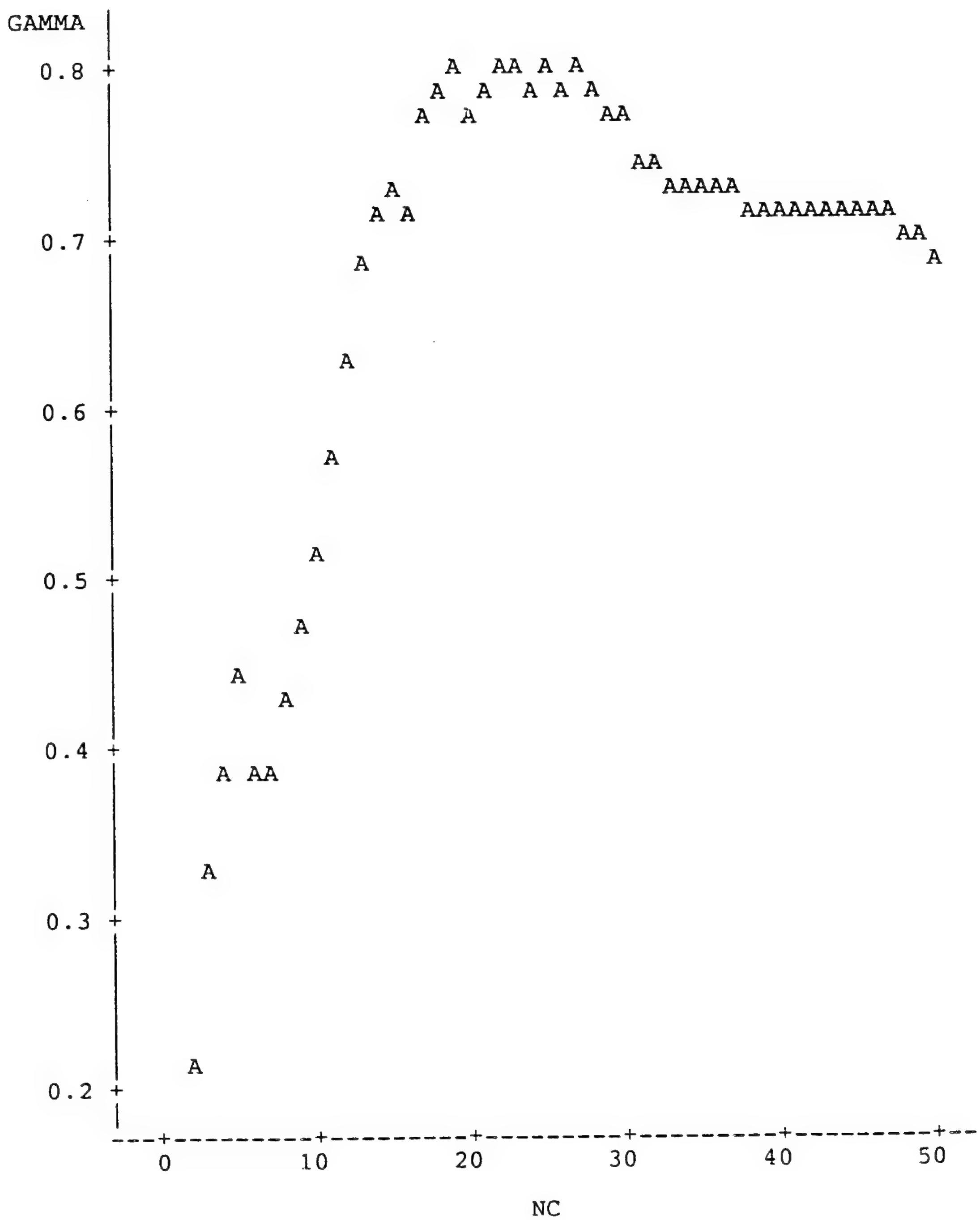


Figure 1. Gamma by Number of Clusters

Table 1.
WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=1 ------------------------------

OBS	MOS	MOSTITLE
1	11H	Heavy Anti-Armor Wpns Inftryman
2	12F	Engineer Tracked Veh Crewman
3	13B	Cannon Crewman
4	13M	MLRS Crewmember
5	15D	Lance Missile Crewmember
6	15E	Pershing Missile Crewmember
7	16B	HERCULES Missile Crewmember
8	16C	HERCULES Fire Control Crewmbr
9	16D	HAWK Missile Crewmember
10	16E	HAWK Fire Control Crewmember
11	16L	Sgt York Air Def Gun Crwmbr
12	16P	ADA Short Range Gunnery Crew
13	16R	ADA Short Range Gunry Crwmbr
14	16S	MANPADS Crewman
15	16T	Patriot Missile Crewmember
16	19E	M48-M60 Armor Crewman
17	19K	M1 ABRAMS Armor Crewman

------------------------------ CLSNUM=2 ------------------------------

OBS	MOS	MOSTITLE
18	61B	Watercraft Operator
19	62E	Hvy Construction Equip Op
20	62F	Crane Operator
21	62G	Quarrying Specialist
22	62H	Concrete & Asphalt Equip Op
23	62J	General Construc Equip Op
24	64C	Motor Transport Operator

------------------------------ CLSNUM=3 ------------------------------

OBS	MOS	MOSTITLE
25	67G	Utility Airplane Repairer
26	67H	Observation Airplane Repairer
27	67N	Utility Helicopter Repairer
28	67R	AH-64 Attack Helicoptr Repr
29	67S	Scout Helicopter Repairer
30	67T	Tact Transp Helicoptr Repr
31	67U	Medium Helicopter Repairer
32	67V	Observ/Scout Helicoptr Repr
33	67Y	AH-1 Attack Helicoptr Repr
34	68B	Aircraft Powerplant Repairer
35	68D	Aircraft Powertrain Repairer
36	68G	Aircraft Structural Repairer
37	68H	Aircraft Pneudraulic Repairer

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=4 ------------------------------

OBS	MOS	MOSTITLE
38	45B	Small Arms Repairer
39	45D	SP Field Artlry Turret Mech
40	45E	M1 ABRAMS Tank Turret Mech
41	45K	Tank Turret Repairer
42	45L	Artillery Repairer
43	45N	M60A1/A3 Tank Turret Mech
44	45T	BFVS Turret Mechanic
45	52C	Utilities Equipment Repairer
46	52D	Power Generator Equip Repr
47	52F	Turbine Engine Generator Repr
48	61C	Watercraft Engineer
49	62B	Construction Equipment Repr
50	63B	Light Wheel Vehicle Mechanic
51	63D	SP Field Artilry System Mech
52	63E	M1 Abrams Tank System Mech
53	63G	Fuel & Elec System Repairer
54	63H	Track Vehicle Repairer
55	63J	Quart&Chem Equipment Repairer
56	63N	M60A1/A3 Tank System Mechanic
57	63S	Heavy Wheel Vehicle Mechanic
58	63T	Bradley System Mechanic
59	63W	Wheel Vehicle Repairer
60	63Y	Track Vehicle Mechanic

------------------------------ CLSNUM=5 ------------------------------

OBS	MOS	MOSTITLE
61	43E	Parachute Rigger
62	55B	Ammunition Specialist
63	55R	Ammo Stock Control&Acct Spec
64	57H	Cargo Specialist
65	76C	Equip Records & Parts Spec
66	76J	Medical Supply Specialist
67	76P	Materl Centrl & Acctng Spec
68	76V	Mat Storage & Handlng Spec
69	76W	Petroleum Supply Specialist
70	76X	Subsistence Supply Specialist
71	76Y	Unit Supply Specialist

------------------------------ CLSNUM=6 ------------------------------

OBS	MOS	MOSTITLE
72	71C	Exec Administrative Assistant
73	71D	Legal Clerk Specialist
74	71G	Patient Admin Specialist
75	71L	Administrative Specialist
76	71M	Chapel Activities Specialist

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------- CLSNUM=6 -------------------------------
(continued)

OBS	MOS	MOSTITLE
77	73C	Finance Specialist
78	73D	Accounting Specialist
79	75B	Personnel Admin Specialist
80	75C	Personnel Management Spec
81	75D	Personnel Records Specialist
82	75E	Personnel Action Specialist
83	75F	Personnel Info Mangmt Spec

------------------------------- CLSNUM=7 -------------------------------

OBS	MOS	MOSTITLE
84	00B	Diver
85	03C	Physical Activities Spec
86	12E	Atomic Demo Munitions Spec
87	43M	Fabric Repair Specialist
88	51G	Materials Quality Specialist
89	51M	Firefighter
90	51N	Water Treatment Specialist
91	52G	Transmisson & Distbution Spec
92	54C	Smoke Operation Specialist
93	54E	NBC Specialist
94	55D	Explsve Ordnance Disposl Spec
95	55G	Nuclear Weapons Maint Spec
96	57E	Laundry & Bath Specialist
97	57F	Graves Registration Spec
98	71N	Traffic Management Coordntor
99	81Q	Terrain Analyst
100	82B	Construction Surveyor
101	82D	Topographic Surveyor
102	91R	Veterinary Food Inspec Spec
103	92C	Petroleum Laboratory Spec
104	92D	Chemical Laboratory Spec
105	93E	Meteorological Observer
106	93F	Field Artlry Meteorlogic Spec
107	94B	Food Service Specialist
108	94F	Hospital Food Service Spec
109	95B	Military Police
110	95C	Correctional Specialist
111	95D	Special Agent

------------------------------- CLSNUM=8 -------------------------------

OBS	MOS	MOSTITLE
112	42C	Orthotic Specialist
113	42D	Dental Laboratory Specialist
114	42E	Optical Laboratory Spec

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=8 ------------------------------
(continued)

OBS	MOS	MOSTITLE
115	91A	Medical Specialist
116	91D	Operating Room Specialist
117	91E	Dental Specialist
118	91F	Psychiatric Specialist
119	91G	Behavioral Science Specialist
120	91H	Orthopedic Specialist
121	91J	Physical Therapy Specialist
122	91L	Occupational Therapy Spec
123	91N	Cardiac Specialist
124	91P	X-Ray Specialist
125	91Q	Pharmacy Specialist
126	91S	Environmental Health Spec
127	91T	Animal Care Specialist
128	91U	Ear, Nose & Throat Specialist
129	91V	Respiratory Specialist
130	91Y	Eye Specialist
131	92B	Medical Laboratory Specialist

------------------------------ CLSNUM=9 ------------------------------

OBS	MOS	MOSTITLE
132	11B	Infantryman
133	11C	Indirect Fire Infantryman
134	11M	Fighting Vehicle Infantryman
135	12B	Combat Engineer
136	12C	Bridge Crewman
137	13F	Fire Support Specialist
138	17C	Field Artlry Target Acq Spec
139	19D	Cavalry Scout
140	82C	Field Artillery Surveyor

------------------------------ CLSNUM=10 ------------------------------

OBS	MOS	MOSTITLE
141	05D	EW/SIGINT Emitter ID/Locator
142	05H	EW/SIGINT Intercept-IMC
143	05K	EW/SIGINT N-M Interceptor
144	16H	ADA Opertns-Intellignce Assis
145	96B	Intelligence Analyst
146	96D	Imagery Analyst
147	96F	Psychological Opertns Spec
148	96H	Aerial Intell Spec
149	97B	Counterintelligence Agents
150	97E	Interrogator
151	97G	Signal Security Specialist
152	98C	EW/SIGINT Analyst

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------- CLSNUM=10 -------------------------------
(continued)

OBS	MOS	MOSTITLE
153	98G	EW/SIGINT Voice Interceptor
154	98J	EW/SIGINT NoncommoIntercept

------------------------------- CLSNUM=11 -------------------------------

OBS	MOS	MOSTITLE
155	21L	PERSHING Electronics Repairer
156	22L	NIKE Test Equipment Repairer
157	22N	NIKEHER MissileLauncherRepair
158	23N	NIKE Track Radar Repairer
159	23U	NIKE Radar-Simulator Repairer
160	24C	Improvd HAWK Firng Sect Mech
161	24E	Improvd HAWK Fire Contrl Mech
162	24G	Imp HAWK Inform CoorCentMech
163	24H	Improvd HAWK Fire Contrl Repr
164	24J	Improved HAWK Pulse Radar Rep
165	24L	ImpHAWK Launch&Mech Sys Repr
166	24M	VULCAN Systems Mechanic
167	24N	CHAPARRAL System Mechanic
168	24Q	NIKE-HERCULES Fire Contrl Mec
169	24U	HERCULES Electronic Mechanic
170	24W	Sgt York Air Def Gun Syst Mec
171	27E	TOW/DRAGON Repairer
172	27F	VULCAN Repairer
173	27G	CHAPARRAL/REDEYE Repairer
174	27L	LANCE System Repairer
175	27M	MLRS Repairer
176	27P	SgtYork Radar/Electron Repr
177	46N	PERSHING ElecMechcal Repairer

------------------------------- CLSNUM=12 -------------------------------

OBS	MOS	MOSTITLE
178	26Q	Tact Satell/Microwave Syst Op
179	26R	Strategic Microwave Syst Op
180	31C	Single Channel Radio Operator
181	31K	Combat Signaler
182	31M	Multichannel Commo Equip Op
183	31N	Tactical Circuit Controller
184	31V	Tactical Commo Syst Op/Mech
185	32D	Station Technical Controller
186	36C	Wire System Instll/Operator
187	36M	Wire Systems Operator
188	72E	Combat Telecomm Ctr Operator
189	72G	Auto Data Telecomm Ctr Oprtor
190	72H	Central Office Operations Op

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

---------------------------- CLSNUM=12 ----------------------------
(continued)

OBS	MOS	MOSTITLE
191	93H	Air Traffic Control Tower Op
192	93J	Air Traff Cntrl Radar Contlr
193	93P	Flight Operations Coordinator

---------------------------- CLSNUM=13 ----------------------------

OBS	MOS	MOSTITLE
194	71Q	Journalist
195	71R	Broadcast Journalist
196	81B	Technical Drafting Specialist
197	81C	Cartographer
198	81E	Illustrator
199	83E	Photo & Layout Specialist
200	83F	Photolithographer
201	84B	Still Photographic Specialist
202	84C	Motion Picture Specialist
203	84F	Audio/TV Specialist

---------------------------- CLSNUM=14 ----------------------------

OBS	MOS	MOSTITLE
204	24K	ImpHAWK Cont-Wave Radar Repr
205	24P	Defense Acq Radar Mechanic
206	26C	Tgt Acq/Surveillance Rdar Rep
207	26D	Ground Cntrl Approch Rdar Rep
208	26E	Aerial Surv Sensor Repairer
209	26F	Aerial Photo-Activ Sensor Rep
210	26H	Air Defense Radar Repairer
211	26K	AerialEWarning/DefEquipRepair
212	26L	Tactical Microwave Syst Repr
213	26T	Radio/TV Systems Specialist
214	26V	Strategic Microwave Syst Repr
215	26Y	SATCOM Equipment Repairer
216	27N	Forwrd Area Alerting Rdar Rep
217	31E	Field Radio Repairer
218	31J	Teletypewriter Repairer
219	31S	Field General Comsec Repairer
220	31T	Field Systems Comsec Repairer
221	32F	Fixed Ciphony Repairer
222	32G	Fixed Crypto Equip Repairer
223	32H	Fixed Station Radio Repairer
224	33P	Ele War/Int StrRecv Subt Repr
225	33Q	Ele War/Int StratProcess&Stor
226	33R	EW/Intrcpt AvSystems Repairer
227	33T	Ele War/Int Tact Systems Rep
228	34T	Tact Computer Systems Repr

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------- CLSNUM=14 -------------------------------
(continued)

OBS	MOS	MOSTITLE
229	35C	Automatic Test Equip Repairer
230	35E	Special Elec Devices Repairer
231	35H	Calibration Specialist
232	35L	Avionic Commo Equip Repairer
233	35M	Avionic Nav/FlightContEq Repr
234	35R	Avionic Special Equip Repr
235	36L	Trans ElectSwitchSys Rep

------------------------------- CLSNUM=15 -------------------------------

OBS	MOS	MOSTITLE
236	44B	Metalworker
237	44E	Machinist
238	51B	Carpentry/Masonry Specialist
239	51C	Structures Specialist
240	51K	Plumber
241	51R	Interior Electrician

------------------------------- CLSNUM=16 -------------------------------

OBS	MOS	MOSTITLE
242	34L	Field Artlry Digital Sys Rep
243	34Y	Field Artlry TactFire Repair
244	35G	Biomedical Equipment Spec
245	35K	Avionic Mechanic
246	36H	Dial/Manual Centrl Office Rep
247	41B	Topographic Instr Rep Spec
248	41C	Fire Contrl Instru Rep Spec
249	41E	Audio/Visual Equip Repairer
250	41J	Office Machine Repairer
251	45G	Fire Control System Repairer
252	68F	Aircraft Electrician
253	68J	Aircraft Fire Control Repr
254	68M	Aircraft Weapon Systems Repr

------------------------------- CLSNUM=17 -------------------------------

OBS	MOS	MOSTITLE
255	13C	TACFIRE Operations Specialist
256	13E	Cannon Fire Direction Speclst
257	13R	Field Artlry Firefindr Rdr Op
258	15J	MLRS/LANCE Ops/FireDir Spec
259	16J	Def Acq Radar Operator
260	17B	Field Artlry Radar Crwmbr

WARDS HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA

SEVENTEEN CLUSTER SOLUTION

------------------------------ CLSNUM=17 ------------------------------
(continued)

OBS	MOS	MOSTITLE
261	21G	PERSHING Elec Materiel Spec
262	24T	PATRIOT Op & System Mechanic
263	25L	ADA Cmnd & Cntrl Syst Op/Repr
264	27B	LandCombat SystemTestSpecial
265	27Q	SgtYork Test Specialist
266	74D	Computer/Machine Operator
267	74F	Programmer Analyst
268	96R	Ground Surv Systems Operator

Appendix F

Average Linkage Hierarchical Cluster Analysis of Project A Data

- Plot of Gamma Values by Number of Clusters
- Composition of Clusters

PLOT OF GAMMA VALUES BY NUMBER OF CLUSTERS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

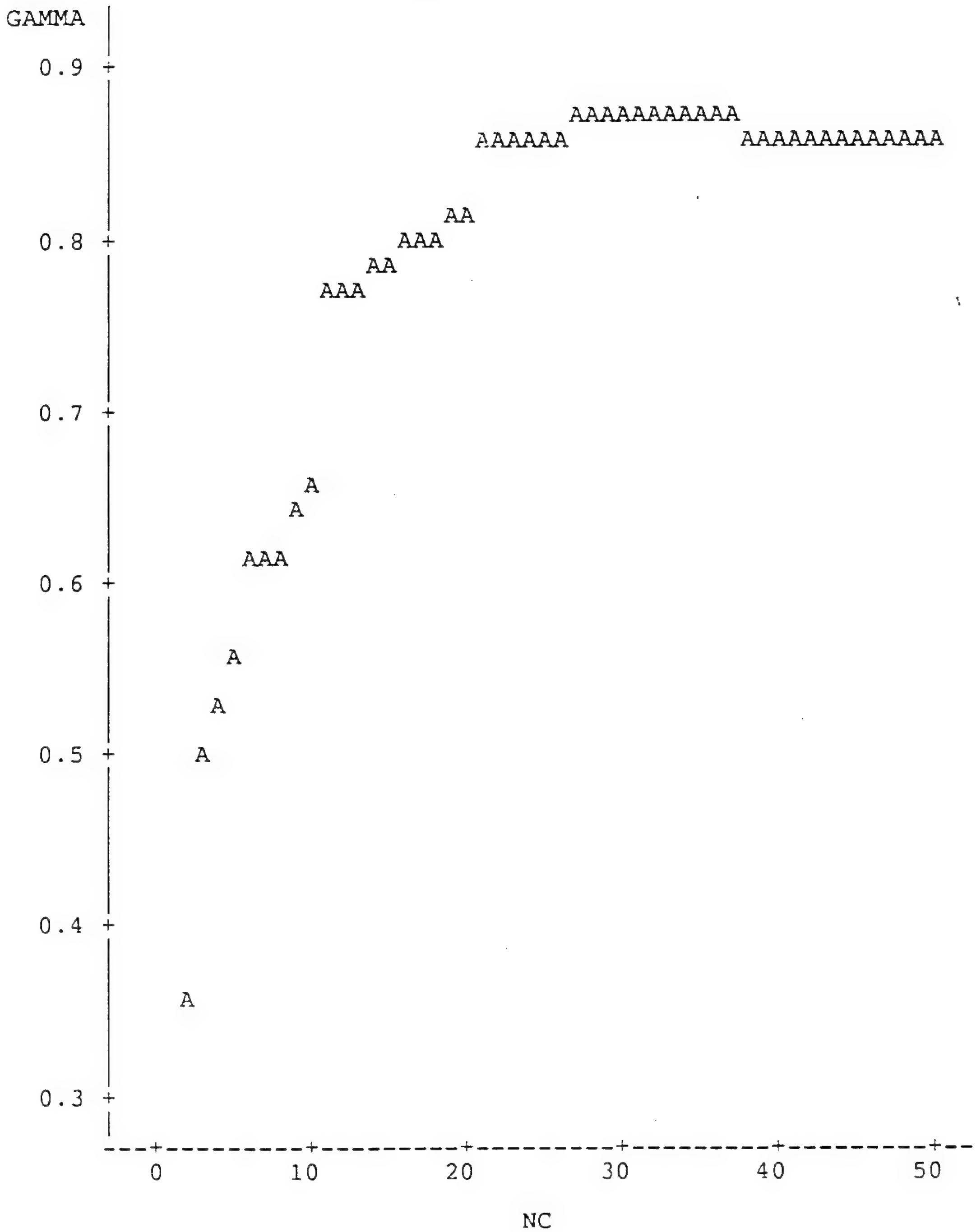


Figure 1. Gamma by Number of Clusters

Table 1.
AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------- CLSNUM=1 --------------------------------

OBS	MOS	MOSTITLE
1	11H	Heavy Anti-Armor Wpns Inftryman
2	12F	Engineer Tracked Veh Crewman
3	13B	Cannon Crewman
4	13E	Cannon Fire Direction Speclst
5	13M	MLRS Crewmember
6	15D	Lance Missile Crewmember
7	15E	Pershing Missile Crewmember
8	15J	MLRS/LANCE Ops/FireDir Spec
9	16B	HERCULES Missile Crewmember
10	16C	HERCULES Fire Control Crewmbr
11	16D	HAWK Missile Crewmember
12	16E	HAWK Fire Control Crewmember
13	16L	Sgt York Air Def Gun Crwmbr
14	16P	ADA Short Range Gunnery Crew
15	16R	ADA Short Range Gunry Crwmbr
16	16S	MANPADS Crewman
17	16T	Patriot Missile Crewmember
18	19E	M48-M60 Armor Crewman
19	19K	M1 ABRAMS Armor Crewman

------------------------------- CLSNUM=2 --------------------------------

OBS	MOS	MOSTITLE
20	54C	Smoke Operation Specialist
21	61B	Watercraft Operator
22	62E	Hvy Construction Equip Op
23	62F	Crane Operator
24	62G	Quarrying Specialist
25	62H	Concrete & Asphalt Equip Op
26	62J	General Construc Equip Op
27	64C	Motor Transport Operator

------------------------------- CLSNUM=3 --------------------------------

OBS	MOS	MOSTITLE
28	67G	Utility Airplane Repairer
29	67H	Observation Airplane Repairer
30	67N	Utility Helicopter Repairer
31	67R	AH-64 Attack Helicoptr Repr
32	67S	Scout Helicopter Repairer
33	67T	Tact Transp Helicoptr Repr
34	67U	Medium Helicopter Repairer
35	67V	Observ/Scout Helicoptr Repr
36	67Y	AH-1 Attack Helicoptr Repr
37	68B	Aircraft Powerplant Repairer
38	68D	Aircraft Powertrain Repairer
39	68G	Aircraft Structural Repairer
40	68H	Aircraft Pneudraulic Repairer
41	68J	Aircraft Fire Control Repr

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------ CLSNUM=3 ------------------------------
(continued)

OBS	MOS	MOSTITLE
42	68M	Aircraft Weapon Systems Repr

------------------------------ CLSNUM=4 ------------------------------

OBS	MOS	MOSTITLE
43	45B	Small Arms Repairer
44	45D	SP Field Artlry Turret Mech
45	45E	M1 ABRAMS Tank Turret Mech
46	45K	Tank Turret Repairer
47	45L	Artillery Repairer
48	45N	M60A1/A3 Tank Turret Mech
49	45T	BFVS Turret Mechanic
50	52C	Utilities Equipment Repairer
51	52D	Power Generator Equip Repr
52	52F	Turbine Engine Generator Repr
53	61C	Watercraft Engineer
54	62B	Construction Equipment Repr
55	63B	Light Wheel Vehicle Mechanic
56	63D	SP Field Artilry System Mech
57	63E	M1 Abrams Tank System Mech
58	63G	Fuel & Elec System Repairer
59	63H	Track Vehicle Repairer
60	63J	Quart&Chem Equipment Repairer
61	63N	M60A1/A3 Tank System Mechanic
62	63S	Heavy Wheel Vehicle Mechanic
63	63T	Bradley System Mechanic
64	63W	Wheel Vehicle Repairer
65	63Y	Track Vehicle Mechanic

------------------------------ CLSNUM=5 ------------------------------

OBS	MOS	MOSTITLE
66	43E	Parachute Rigger
67	55B	Ammunition Specialist
68	55R	Ammo Stock Control&Acct Spec
69	57H	Cargo Specialist
70	71N	Traffic Management Coordntor
71	76C	Equip Records & Parts Spec
72	76J	Medical Supply Specialist
73	76P	Materl Centrl & Acctng Spec
74	76V	Mat Storage & Handlng Spec
75	76W	Petroleum Supply Specialist
76	76X	Subsistence Supply Specialist
77	76Y	Unit Supply Specialist

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

-------------------------------- CLSNUM=6 --------------------------------

OBS	MOS	MOSTITLE
78	03C	Physical Activities Spec
79	57F	Graves Registration Spec
80	71C	Exec Administrative Assistant
81	71D	Legal Clerk Specialist
82	71G	Patient Admin Specialist
83	71L	Administrative Specialist
84	71M	Chapel Activities Specialist
85	73C	Finance Specialist
86	73D	Accounting Specialist
87	75B	Personnel Admin Specialist
88	75C	Personnel Management Spec
89	75D	Personnel Records Specialist
90	75E	Personnel Action Specialist
91	75F	Personnel Info Mangmt Spec

-------------------------------- CLSNUM=7 --------------------------------

OBS	MOS	MOSTITLE
92	91R	Veterinary Food Inspec Spec
93	94B	Food Service Specialist
94	94F	Hospital Food Service Spec

-------------------------------- CLSNUM=8 --------------------------------

OBS	MOS	MOSTITLE
95	42C	Orthotic Specialist
96	42D	Dental Laboratory Specialist
97	42E	Optical Laboratory Spec
98	91A	Medical Specialist
99	91D	Operating Room Specialist
100	91E	Dental Specialist
101	91F	Psychiatric Specialist
102	91G	Behavioral Science Specialist
103	91H	Orthopedic Specialist
104	91J	Physical Therapy Specialist
105	91L	Occupational Therapy Spec
106	91N	Cardiac Specialist
107	91P	X-Ray Specialist
108	91Q	Pharmacy Specialist
109	91S	Environmental Health Spec
110	91T	Animal Care Specialist
111	91U	Ear, Nose & Throat Specialist
112	91V	Respiratory Specialist
113	91Y	Eye Specialist
114	92B	Medical Laboratory Specialist

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------- CLSNUM=9 -------------------------------

OBS	MOS	MOSTITLE
115	11B	Infantryman
116	11C	Indirect Fire Infantryman
117	11M	Fighting Vehicle Infantryman
118	12B	Combat Engineer
119	12C	Bridge Crewman
120	13F	Fire Support Specialist
121	17C	Field Artlry Target Acq Spec
122	19D	Cavalry Scout
123	82C	Field Artillery Surveyor

------------------------------- CLSNUM=10 -------------------------------

OBS	MOS	MOSTITLE
124	05D	EW/SIGINT Emitter ID/Locator
125	05H	EW/SIGINT Intercept-IMC
126	05K	EW/SIGINT N-M Interceptor
127	16H	ADA Opertns-Intellignce Assis
128	96B	Intelligence Analyst
129	96D	Imagery Analyst
130	96F	Psychological Opertns Spec
131	96H	Aerial Intell Spec
132	97B	Counterintelligence Agents
133	97E	Interrogator
134	97G	Signal Security Specialist
135	98C	EW/SIGINT Analyst
136	98G	EW/SIGINT Voice Interceptor
137	98J	EW/SIGINT NoncommoIntercept

------------------------------- CLSNUM=11 -------------------------------

OBS	MOS	MOSTITLE
138	21G	PERSHING Elec Materiel Spec
139	21L	PERSHING Electronics Repairer
140	22L	NIKE Test Equipment Repairer
141	22N	NIKEHER MissileLauncherRepair
142	23N	NIKE Track Radar Repairer
143	23U	NIKE Radar-Simulator Repairer
144	24C	Improvd HAWK Firng Sect Mech
145	24E	Improvd HAWK Fire Contrl Mech
146	24G	Imp HAWK Inform CoorCentMech
147	24H	Improvd HAWK Fire Contrl Repr
148	24J	Improved HAWK Pulse Radar Rep
149	24K	ImpHAWK Cont-Wave Radar Repr
150	24L	ImpHAWK Launch&Mech Sys Repr
151	24M	VULCAN Systems Mechanic
152	24N	CHAPARRAL System Mechanic
153	24Q	NIKE-HERCULES Fire Contrl Mec
154	24T	PATRIOT Op & System Mechanic
155	24U	HERCULES Electronic Mechanic

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------ CLSNUM=11 ------------------------------
(continued)

OBS	MOS	MOSTITLE
156	24W	Sgt York Air Def Gun Syst Mec
157	27B	LandCombat SystemTestSpecial
158	27E	TOW/DRAGON Repairer
159	27F	VULCAN Repairer
160	27G	CHAPARRAL/REDEYE Repairer
161	27L	LANCE System Repairer
162	27M	MLRS Repairer
163	27P	SgtYork Radar/Electron Repr
164	27Q	SgtYork Test Specialist
165	34L	Field Artlry Digital Sys Rep
166	34Y	Field Artlry TactFire Repair
167	41C	Fire Contrl Instru Rep Spec
168	45G	Fire Control System Repairer
169	46N	PERSHING ElecMechcal Repairer

------------------------------ CLSNUM=12 ------------------------------

OBS	MOS	MOSTITLE
170	13R	Field Artlry Firefindr Rdr Op
171	16J	Def Acq Radar Operator
172	17B	Field Artlry Radar Crwmbr
173	25L	ADA Cmnd & Cntrl Syst Op/Repr
174	26Q	Tact Satell/Microwave Syst Op
175	26R	Strategic Microwave Syst Op
176	31C	Single Channel Radio Operator
177	31K	Combat Signaler
178	31M	Multichannel Commo Equip Op
179	31N	Tactical Circuit Controller
180	31V	Tactical Commo Syst Op/Mech
181	32D	Station Technical Controller
182	36C	Wire System Instll/Operator
183	36M	Wire Systems Operator
184	72E	Combat Telecomm Ctr Operator
185	72G	Auto Data Telecomm Ctr Oprtor
186	72H	Central Office Operations Op
187	93H	Air Traffic Control Tower Op
188	93J	Air Traff Cntrl Radar Contlr
189	93P	Flight Operations Coordinator
190	96R	Ground Surv Systems Operator

------------------------------ CLSNUM=13 ------------------------------

OBS	MOS	MOSTITLE
191	71Q	Journalist
192	71R	Broadcast Journalist
193	81B	Technical Drafting Specialist
194	81C	Cartographer
195	81E	Illustrator

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------- CLSNUM=13 -------------------------------
(continued)

OBS	MOS	MOSTITLE
196	83E	Photo & Layout Specialist
197	83F	Photolithographer
198	84B	Still Photographic Specialist
199	84C	Motion Picture Specialist
200	84F	Audio/TV Specialist

------------------------------- CLSNUM=14 -------------------------------

OBS	MOS	MOSTITLE
201	24P	Defense Acq Radar Mechanic
202	26C	Tgt Acq/Surveillance Rdar Rep
203	26D	Ground Cntrl Approch Rdar Rep
204	26E	Aerial Surv Sensor Repairer
205	26F	Aerial Photo-Activ Sensor Rep
206	26H	Air Defense Radar Repairer
207	26K	AerialEWarning/DefEquipRepair
208	26L	Tactical Microwave Syst Repr
209	26T	Radio/TV Systems Specialist
210	26V	Strategic Microwave Syst Repr
211	26Y	SATCOM Equipment Repairer
212	27N	Forwrd Area Alerting Rdar Rep
213	31E	Field Radio Repairer
214	31J	Teletypewriter Repairer
215	31S	Field General Comsec Repairer
216	31T	Field Systems Comsec Repairer
217	32F	Fixed Ciphony Repairer
218	32G	Fixed Crypto Equip Repairer
219	32H	Fixed Station Radio Repairer
220	33P	Ele War/Int StrRecv Subt Repr
221	33Q	Ele War/Int StratProcess&Stor
222	33R	EW/Intrcpt AvSystems Repairer
223	33T	Ele War/Int Tact Systems Rep
224	34T	Tact Computer Systems Repr
225	35C	Automatic Test Equip Repairer
226	35E	Special Elec Devices Repairer
227	35G	Biomedical Equipment Spec
228	35H	Calibration Specialist
229	35K	Avionic Mechanic
230	35L	Avionic Commo Equip Repairer
231	35M	Avionic Nav/FlightContEq Repr
232	35R	Avionic Special Equip Repr
233	36H	Dial/Manual Centrl Office Rep
234	36L	Trans ElectSwitchSys Rep
235	41B	Topographic Instr Rep Spec
236	41E	Audio/Visual Equip Repairer
237	41J	Office Machine Repairer
238	68F	Aircraft Electrician

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------- CLSNUM=15 -------------------------------

OBS	MOS	MOSTITLE
239	43M	Fabric Repair Specialist
240	44B	Metalworker
241	44E	Machinist
242	51B	Carpentry/Masonry Specialist
243	51C	Structures Specialist
244	51K	Plumber
245	51R	Interior Electrician
246	52G	Transmisson & Distbution Spec
247	57E	Laundry & Bath Specialist

------------------------------- CLSNUM=16 -------------------------------

OBS	MOS	MOSTITLE
248	95B	Military Police
249	95C	Correctional Specialist
250	95D	Special Agent

------------------------------- CLSNUM=17 -------------------------------

OBS	MOS	MOSTITLE
251	51G	Materials Quality Specialist
252	51N	Water Treatment Specialist
253	92C	Petroleum Laboratory Spec
254	92D	Chemical Laboratory Spec
255	93E	Meteorological Observer
256	93F	Field Artlry Meteorlogic Spec

------------------------------- CLSNUM=18 -------------------------------

OBS	MOS	MOSTITLE
257	12E	Atomic Demo Munitions Spec
258	54E	NBC Specialist
259	55D	Explsve Ordnance Disposl Spec
260	55G	Nuclear Weapons Maint Spec

------------------------------- CLSNUM=19 -------------------------------

OBS	MOS	MOSTITLE
261	81Q	Terrain Analyst
262	82B	Construction Surveyor
263	82D	Topographic Surveyor

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS OF PROJECT A DATA
TWENTY-ONE CLUSTER SOLUTION

------------------------------- CLSNUM=20 --------------------------------

OBS	MOS	MOSTITLE
264	13C	TACFIRE Operations Specialist
265	74D	Computer/Machine Operator
266	74F	Programmer Analyst

------------------------------- CLSNUM=21 --------------------------------

OBS	MOS	MOSTITLE
267	00B	Diver
268	51M	Firefighter

Appendix G

Ward Hierarchical Cluster Analysis of Synthetic Validity OJI Data

- Plot of Cubic Clustering Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

PLOT OF CUBIC CLUSTERING CRITERION BY NUMBER OF CLUSTERS

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.

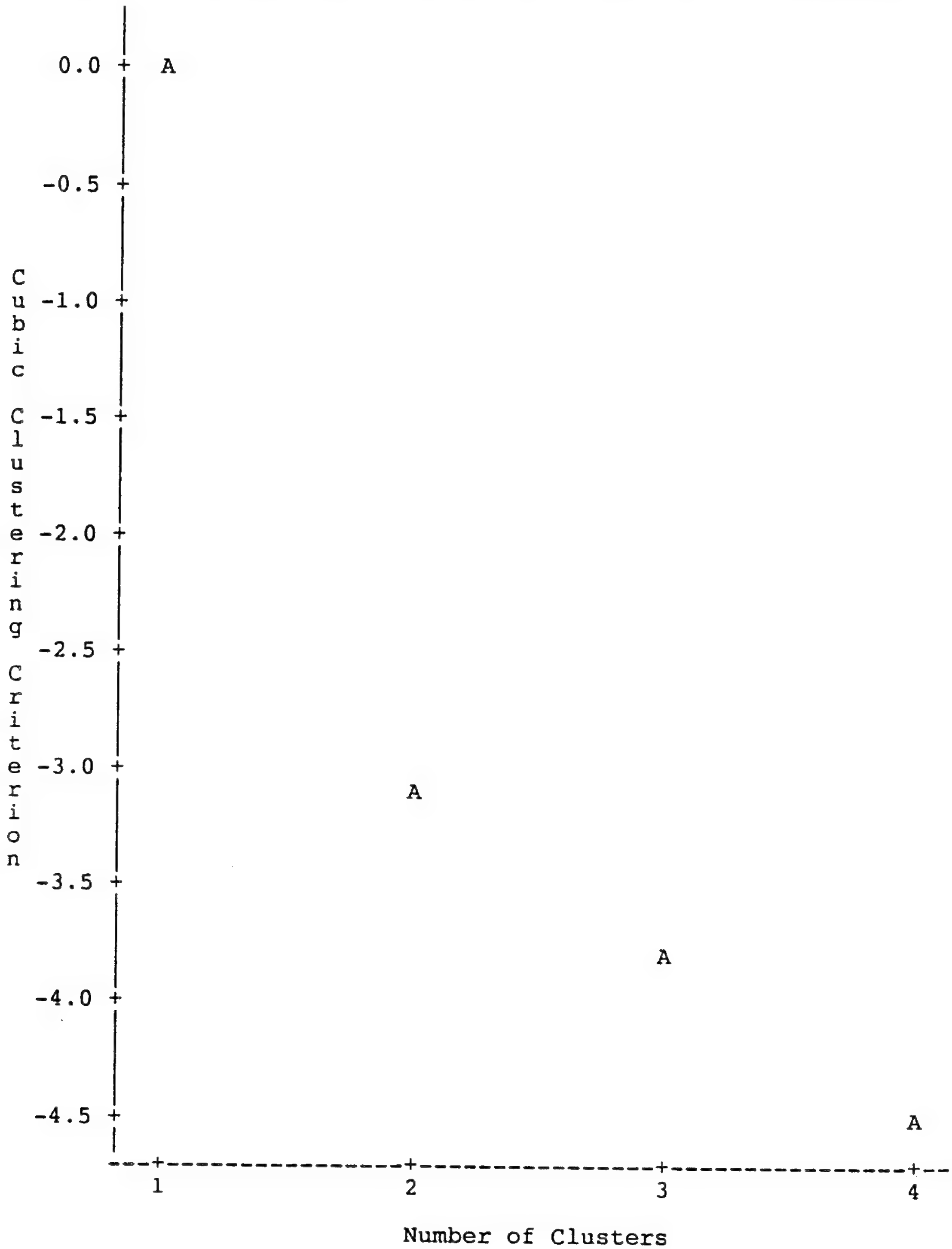


Figure 1. CCC by Number of Clusters

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

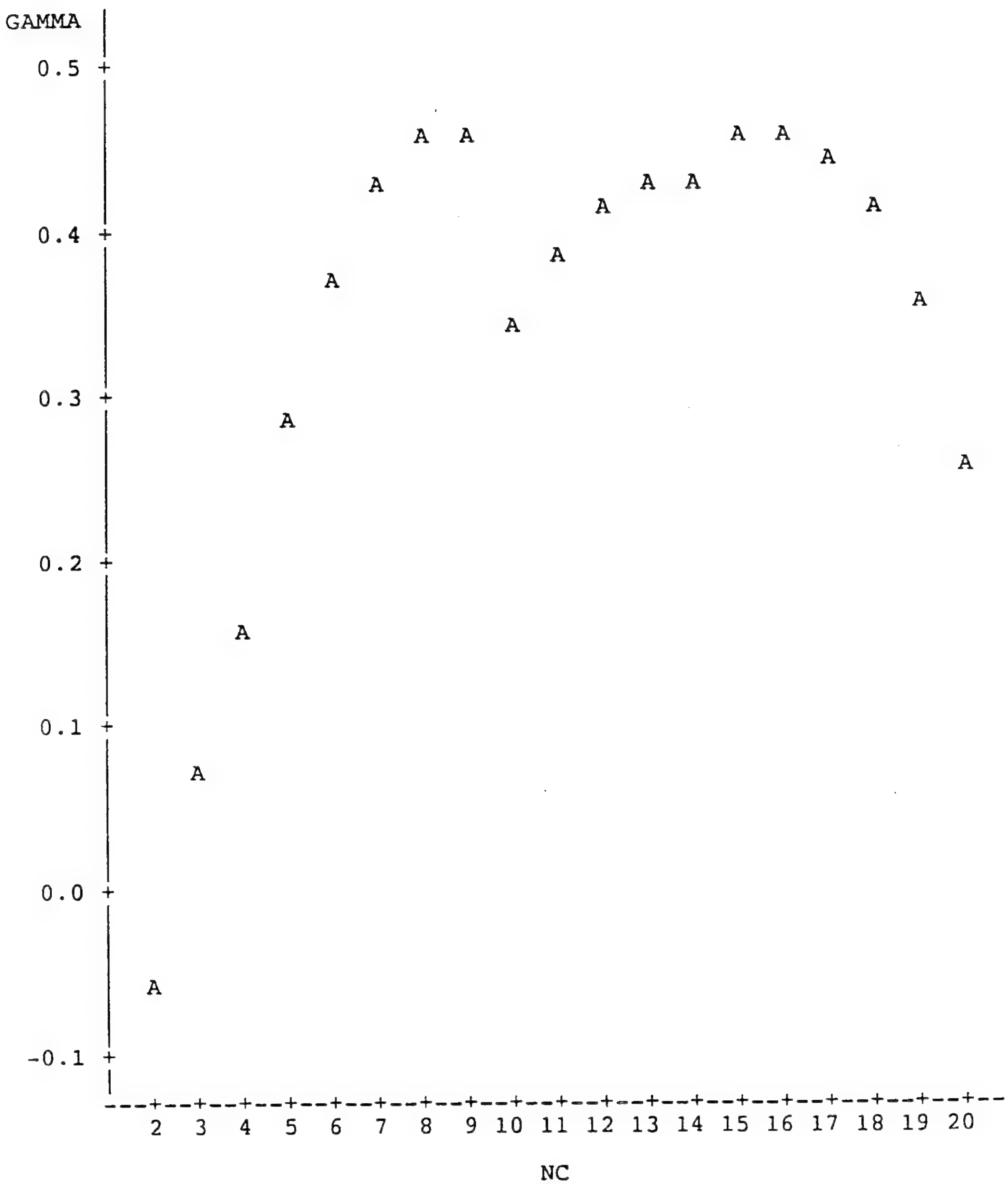


Figure 2. Gamma by Number of Clusters

Table 1
WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
EIGHT CLUSTER SOLUTION

---------------------------------- CLSNUM=1 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
1	11B	Infantryman	2.92344
2	16S	MANPADS Crewman	2.82181
3	19K	M1 ABRAMS Armor Crewman	3.58506
4	27E	TOW/DRAGON Repairer	3.56884
5	29E	Radio Repairer	3.68931
6	63B	Light Wheel Vehicle Mechanic	2.65875
7	67N	Utility Helicopter Repairer	3.08387
8	88M	Motor Transport Operator	3.51946

---------------------------------- CLSNUM=2 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
9	31C	Single Channel Radio Operator	2.52782
10	54B	Chemical Operations Specialist	2.43594
11	94B	Food Service Specialist	3.59355

---------------------------------- CLSNUM=3 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
12	12B	Combat Engineer	2.85401
13	76Y	Unit Supply Specialist	2.85401

---------------------------------- CLSNUM=4 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
14	31D	Mobile Subscriber Equipment	3.48259
15	51B	Carpentry/Masonry Specialist	3.56486
16	71L	Administrative Specialist	3.43608

---------------------------------- CLSNUM=5 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
17	13B	Cannon Crewman	3.08094
18	55B	Ammunition Specialist	3.08094

---------------------------------- CLSNUM=6 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
19	96B	Intelligence Analyst	0

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
EIGHT CLUSTER SOLUTION

------------------------------ CLSNUM=7 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
20	91A	Medical Specialist	0

------------------------------ CLSNUM=8 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
21	95B	Military Police	0

Table 2.
WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
EIGHT CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=1

N	Mean	Std Dev	Minimum	Maximum
8	3.2313172	0.4042444	2.6587493	3.6893092

CLSNUM=2

N	Mean	Std Dev	Minimum	Maximum
3	2.8524381	0.6434644	2.4359410	3.5935507

CLSNUM=3

N	Mean	Std Dev	Minimum	Maximum
2	2.8540053	0	2.8540053	2.8540053

CLSNUM=4

N	Mean	Std Dev	Minimum	Maximum
3	3.4945098	0.0652136	3.4360757	3.5648597

CLSNUM=5

N	Mean	Std Dev	Minimum	Maximum
2	3.0809428	0	3.0809428	3.0809428

CLSNUM=6

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
EIGHT CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

---------------------------------- CLSNUM=7 ----------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

---------------------------------- CLSNUM=8 ----------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3.

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
EIGHT CLUSTER SOLUTION
MEAN CLUSTER FACTOR SCORES

OBS	CLSNUM	FREQ	LGEDIST	PRIN1	PRIN2	PRIN3	PRIN4
1	1	8	3.68931	0.16195	-0.37306	-0.19963	0.43834
2	2	3	3.59355	-0.04545	-0.43279	0.34426	-0.14932
3	3	2	2.85401	0.01871	-0.08718	-0.43833	-0.67815
4	4	3	3.56486	-0.61967	-0.20178	0.36720	0.19242
5	5	2	3.08094	0.33972	-0.10698	-0.74740	-0.28323
6	6	1	0.00000	-0.04311	0.55400	1.65470	-2.56926
7	7	1	0.00000	-0.27141	0.70482	-0.13430	-0.50050
8	8	1	0.00000	0.29742	4.01766	0.31370	1.35647

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
1	0.58292	-0.57820	0.04890	-0.14441	-0.37943	-0.11787
2	-0.18999	0.32584	-0.54058	0.32798	-0.65099	0.19024
3	-0.75270	0.90145	0.74362	-0.45173	-0.49458	-0.10870
4	-1.11970	-0.33599	-0.34902	-0.41417	1.42858	-0.79440
5	0.47559	1.98647	-0.04663	-0.37203	0.55227	0.72827
6	-0.47601	-1.37828	0.47188	-0.39314	0.13374	1.76364
7	0.12441	0.22752	0.72054	4.03065	0.55734	-0.50364
8	0.17154	0.03094	-0.30882	-0.57616	-0.10378	0.25632

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15	PRIN16
1	-0.15752	-0.47769	-0.08594	-0.00239	0.00331	0.03206
2	1.44209	1.06595	0.95494	0.22960	-0.44511	-0.01565
3	-0.10685	0.36469	-0.90716	1.03247	1.64277	-1.05295
4	-0.22388	0.21539	-0.43981	-0.07177	-0.61796	-0.12835
5	-1.07278	0.15683	0.38772	-0.90186	-0.48833	0.79498
6	0.07216	-0.82405	0.24126	-1.25092	0.74909	0.62108
7	-0.33510	-0.49346	-0.29432	0.36233	0.02985	0.32632
8	0.22770	0.25194	0.23404	0.17302	0.07488	-0.25594

Appendix H

Average Linkage Hierarchical Cluster Analysis of Synthetic Validity OJI Data

- Plot of Cubic Clustering Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.

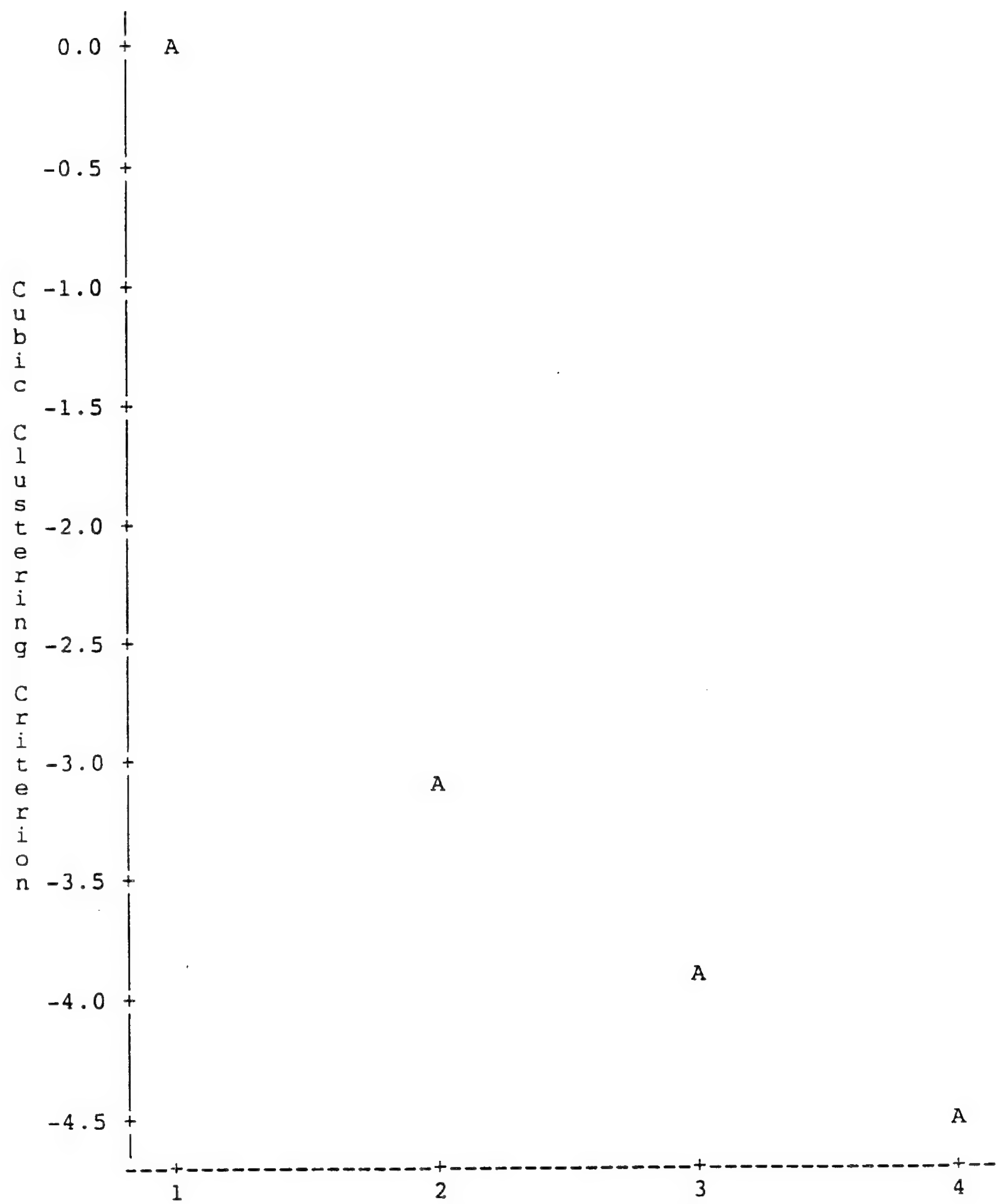


Figure 1. CCC by Number of Clusters

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

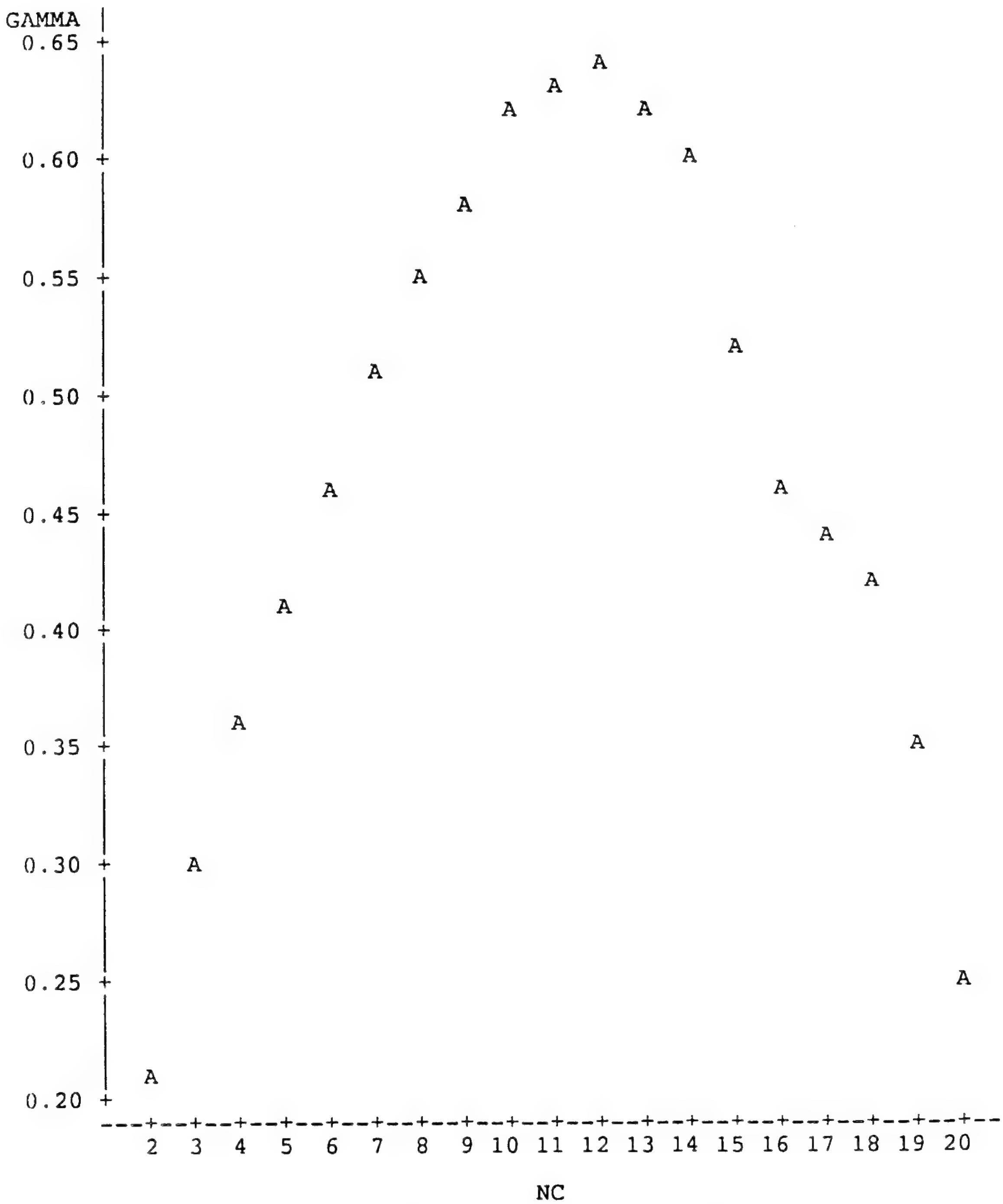


Figure 2. Gamma by Number of Clusters

Table 1
AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION

---------------------------------- CLSNUM=1 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
1	11B	Infantryman	2.99355
2	12B	Combat Engineer	3.96411
3	16S	MANPADS Crewman	2.91997
4	19K	M1 ABRAMS Armor Crewman	3.77634
5	27E	TOW/DRAGON Repairer	3.76437
6	29E	Radio Repairer	3.78811
7	31C	Single Channel Radio Operator	3.41772
8	31D	Mobile Subscriber Equipment	4.05456
9	54B	Chemical Operations Specialist	3.25475
10	63B	Light Wheel Vehicle Mechanic	2.83969
11	67N	Utility Helicopter Repairer	3.09279
12	88M	Motor Transport Operator	3.63563

---------------------------------- CLSNUM=2 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
13	76Y	Unit Supply Specialist	0

---------------------------------- CLSNUM=3 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
14	71L	Administrative Specialist	0

---------------------------------- CLSNUM=4 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
15	55B	Ammunition Specialist	0

---------------------------------- CLSNUM=5 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
16	51B	Carpentry/Masonry Specialist	0

---------------------------------- CLSNUM=6 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
17	13B	Cannon Crewman	0

AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION

------------------------------- CLSNUM=7 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
18	96B	Intelligence Analyst	0

------------------------------- CLSNUM=8 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
19	94B	Food Service Specialist	0

------------------------------- CLSNUM=9 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
20	91A	Medical Specialist	0

------------------------------- CLSNUM=10 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
21	95B	Military Police	0

Table 2.
AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=1

N	Mean	Std Dev	Minimum	Maximum
12	3.4584665	0.4276312	2.8396881	4.0545550

CLSNUM=2

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=3

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=4

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=5

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=6

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=7

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=8

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=9

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=10

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3.
AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION
MEAN CLUSTER PRINCIPAL COMPONENT SCORES

OBS	CLSNUM	FREQ	LGEDIST	PRIN1	PRIN2	PRIN3	PRIN4
1	1	12	4.05456	0.33819	-0.38951	0.16881	0.25323
2	2	1	0.00000	-1.72717	-0.09264	-0.56586	-0.66921
3	3	1	0.00000	-2.02128	0.41597	-0.09448	-1.38610
4	4	1	0.00000	-0.67466	0.03274	-0.79051	-0.22871
5	5	1	0.00000	0.48131	-0.40400	-0.87607	1.15929
6	6	1	0.00000	1.35409	-0.24669	-0.70428	-0.33774
7	7	1	0.00000	-0.04311	0.55400	1.65470	-2.56926
8	8	1	0.00000	-1.45349	-0.30776	-0.82867	0.13695
9	9	1	0.00000	-0.27141	0.70482	-0.13430	-0.50050
10	10	1	0.00000	0.29742	4.01766	0.31370	1.35647

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
1	0.19948	-0.27206	-0.15120	-0.05502	-0.36989	-0.15856
2	0.24122	0.96394	0.74618	-0.63031	0.23631	1.00246
3	-0.02133	-0.59641	0.58385	-0.94271	0.88692	-2.30650
4	-0.56347	1.97140	-0.29016	-0.52110	-0.89249	0.89679
5	-3.04245	-0.76095	0.62027	-0.24751	2.00416	0.28422
6	1.51465	2.00153	0.19689	-0.22295	1.99702	0.55974
7	-0.47601	-1.37828	0.47188	-0.39314	0.13374	1.76364
8	-0.34233	0.80503	-0.92621	0.16350	-0.38058	-0.05036
9	0.12441	0.22752	0.72054	4.03065	0.55734	-0.50364
10	0.17154	0.03094	-0.30882	-0.57616	-0.10378	0.25632

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15	PRIN16
1	-0.06162	0.04820	-0.14813	0.09984	0.01116	-0.08222
2	-0.17610	0.14298	-0.47824	2.27991	1.57933	-1.75315
3	-0.34338	0.88374	-0.23389	-0.52184	-1.62093	-0.32850
4	-2.04973	-0.12379	0.41826	0.26454	-1.00536	2.37607
5	0.00911	-0.50982	0.95462	0.24195	-0.52744	-0.18999
6	-0.09582	0.43744	0.35718	-2.06825	0.02871	-0.78612
7	0.07216	-0.82405	0.24126	-1.25092	0.74909	0.62108
8	3.43055	-0.34337	0.57864	-0.67882	0.55796	0.97680
9	-0.33510	-0.49346	-0.29432	0.36233	0.02985	0.32632
10	0.22770	0.25194	0.23404	0.17302	0.07488	-0.25594

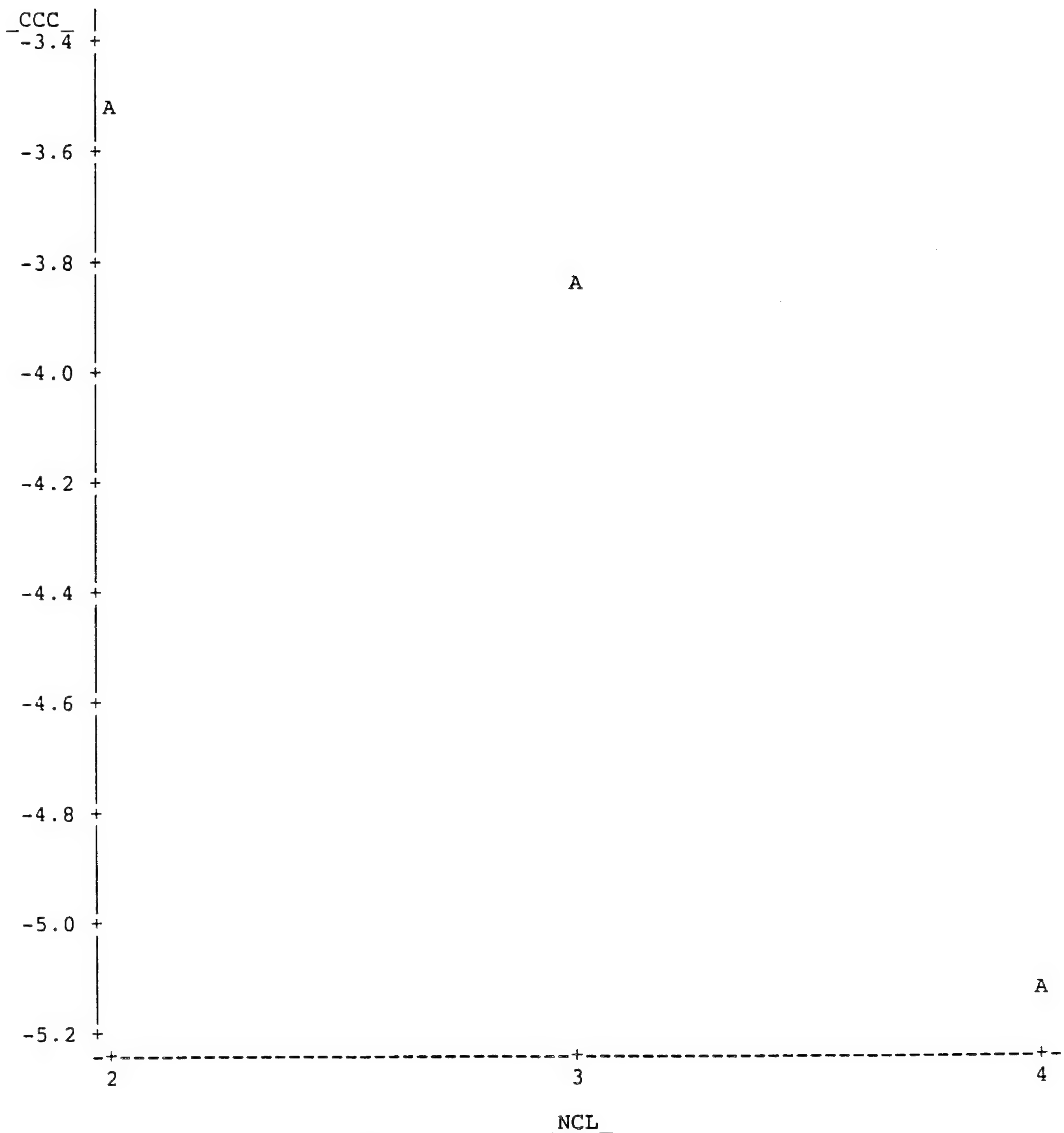
Appendix I

K-Means Cluster Analysis of Synthetic Validity OJI Data

- Plot of Cubic Clustering Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

SAS MODIFIED K-MEANS ITERATIVE CLUSTER ANALYSIS
CCC DIFFERENCES

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.



Gigure 1. CCC by Number of Clusters

SAS MODIFIED K-MEANS ITERATIVE CLUSTER ANALYSIS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

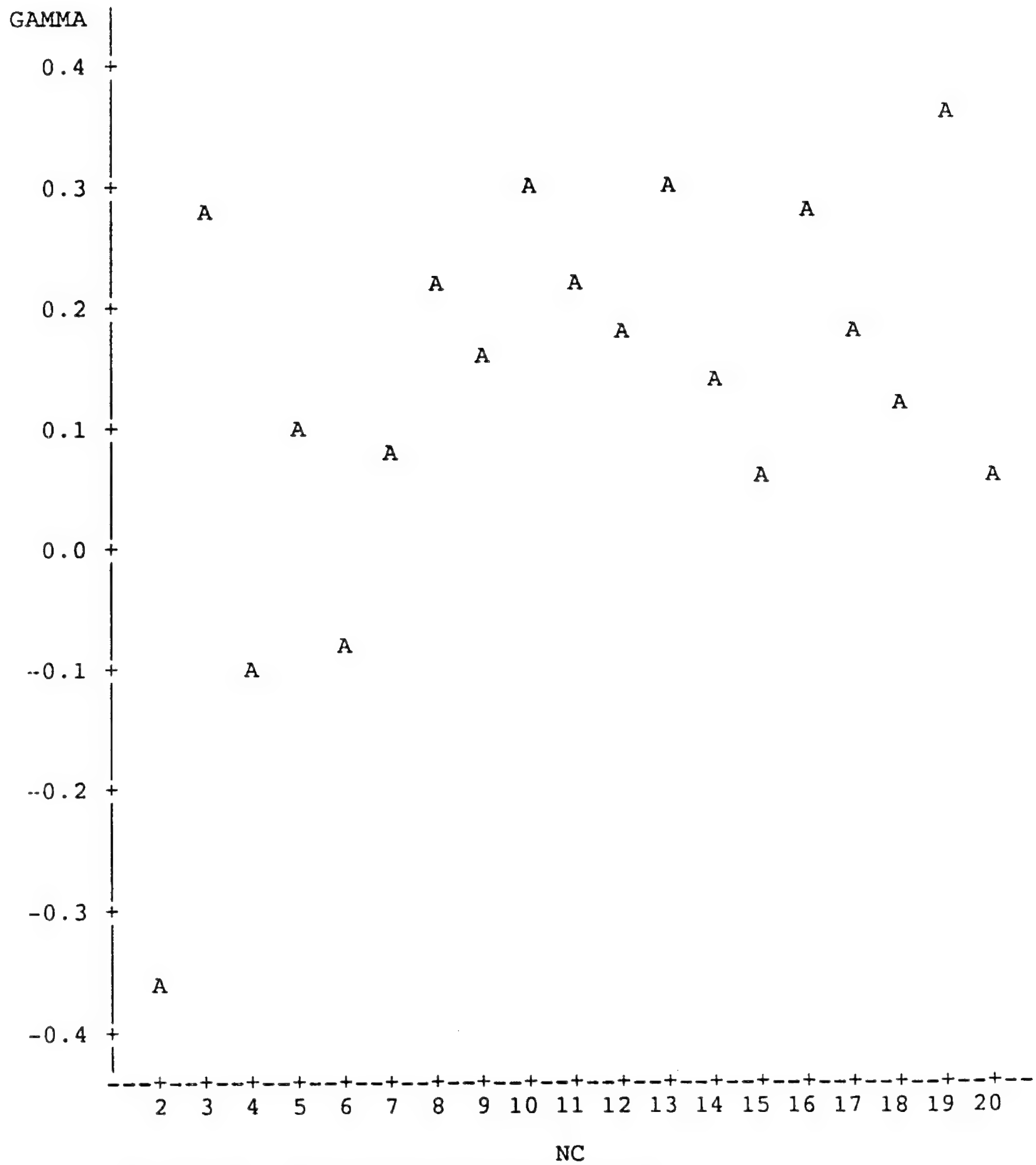


Figure 2. Gamma by Number of Clusters

Table 1
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION

------------------------------ Cluster=1 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
1	94B	Food Service Specialist	0

------------------------------ Cluster=2 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
2	63B	Light Wheel Vehicle Mechanic	0

------------------------------ Cluster=3 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
3	88M	Motor Transport Operator	0

------------------------------ Cluster=4 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
4	11B	Infantryman	2.94358
5	16S	MANPADS Crewman	3.08304
6	19K	M1 ABRAMS Armor Crewman	3.73339
7	29E	Radio Repairer	3.86228
8	12B	Combat Engineer	3.90373
9	31D	Mobile Subscriber Equipment	4.07801
10	71L	Administrative Specialist	4.13386
11	76Y	Unit Supply Specialist	4.14541
12	55B	Ammunition Specialist	4.19715
13	51B	Carpentry/Masonry Specialist	4.21129
14	13B	Cannon Crewman	4.22277
15	95B	Military Police	4.26833

------------------------------ Cluster=5 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
16	96B	Intelligence Analyst	0

------------------------------ Cluster=6 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
17	27E	TOW/DRAGON Repairer	0

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION

------------------------------- Cluster=7 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
18	91A	Medical Specialist	0

------------------------------- Cluster=8 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
19	31C	Single Channel Radio Operator	0

------------------------------- Cluster=9 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE
20	67N	Utility Helicopter Repairer	0

------------------------------- Cluster=10 -------------------------------

OBS	MOS	MOSTITLE	DISTANCE

Table 2
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

Cluster=1

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=2

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=3

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=4

N	Mean	Std Dev	Minimum	Maximum
12	3.8985686	0.4454806	2.9435756	4.2683254

Cluster=5

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=6

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

Cluster=7

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=8

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=9

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=10

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3.
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY OJI DATA
TEN CLUSTER SOLUTION
MEAN CLUSTER PRINCIPAL COMPONENT SCORES

OBS	FREQ	LGEDIST	NEAREST	PRIN1	PRIN2	PRIN3	PRIN4
1	1	0.00000	4	-1.45349	-0.30776	-0.82867	0.13695
2	1	0.00000	9	-0.35237	-0.39903	-0.74343	1.18719
3	1	0.00000	2	-0.68529	-0.32751	-1.31049	0.29866
4	12	4.26833	2	0.14672	0.16703	-0.15851	-0.03067
5	1	0.00000	4	-0.04311	0.55400	1.65470	-2.56926
6	1	0.00000	4	0.06958	-0.68118	1.10102	1.14129
7	1	0.00000	4	-0.27141	0.70482	-0.13430	-0.50050
8	1	0.00000	10	0.58823	-0.61715	1.70462	-0.00107
9	1	0.00000	2	-0.34172	-0.55708	0.30183	1.25862
10	1	0.00000	8	0.72890	-0.37345	0.15684	-0.58383

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10	PRIN11
1	-0.34233	0.80503	-0.92621	0.16350	-0.38058	-0.05036	3.43055
2	0.46256	-1.61588	0.63775	0.31260	-0.81999	-0.00851	-0.62641
3	0.06086	-0.75626	-1.87959	0.45511	-0.89226	0.64058	-1.03227
4	-0.06439	0.21604	-0.11046	-0.41303	0.31122	-0.32525	-0.27050
5	-0.47601	-1.37828	0.47188	-0.39314	0.13374	1.76364	0.07216
6	0.84554	0.46725	2.02812	-0.38311	-0.40622	0.29944	0.61493
7	0.12441	0.22752	0.72054	4.03065	0.55734	-0.50364	-0.33510
8	0.02998	0.41274	-0.50260	0.49143	-0.17604	-0.24771	0.23894
9	0.32534	-0.51443	0.96856	-0.04972	-0.35425	1.14074	0.22637
10	-0.25762	-0.24024	-0.19294	0.32902	-1.39635	0.86878	0.65679

OBS	PRIN12	PRIN13	PRIN14	PRIN15	PRIN16
1	-0.34337	0.57864	-0.67882	0.55796	0.97680
2	1.10322	-0.26568	-1.09865	0.37806	0.00590
3	-0.51961	0.13608	-1.10479	0.05167	-1.74510
4	-0.11420	-0.05039	0.20798	0.14036	0.03794
5	-0.82405	0.24126	-1.25092	0.74909	0.62108
6	-1.68962	-0.29945	0.18506	-1.45998	-0.88927
7	-0.49346	-0.29432	0.36233	0.02985	0.32632
8	1.62805	1.90414	1.18632	-0.43266	-0.03417
9	0.59605	-1.77801	-0.27759	-0.09764	1.27268
10	1.91317	0.38203	0.18131	-1.46064	-0.98958

Appendix J

Ward Hierarchical Cluster Analysis of Synthetic Validity CTI Data

- Plot of Cubic Clustering Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

WARDS HIERARCHICAL CLUSTER ANALYSIS

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.

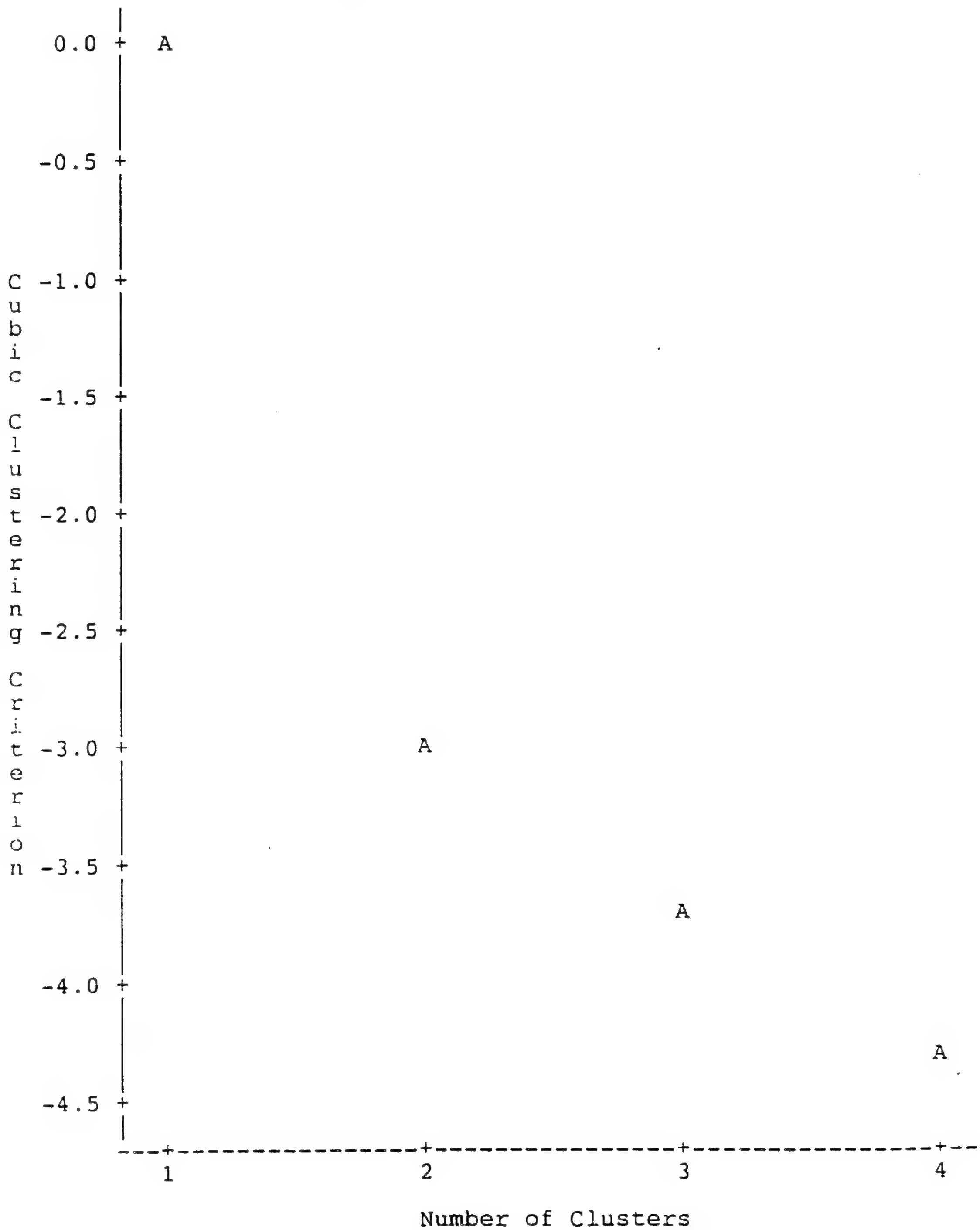


Figure 1. CCC by Number of Clusters

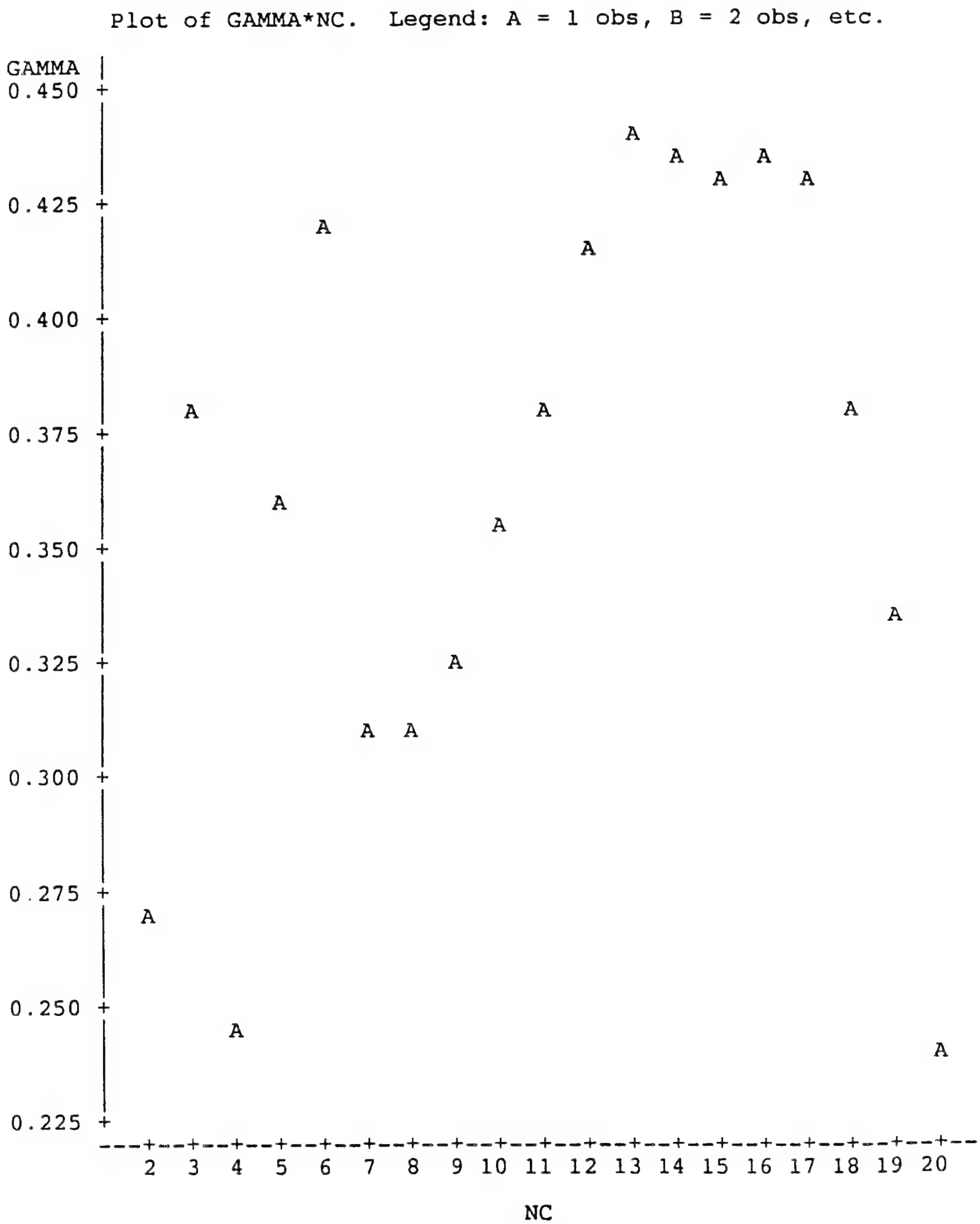


Figure 2. Gamma by Number of Clusters

Table 2.
WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
THIRTEEN CLUSTER SOLUTION

CLSNUM=1

OBS	MOS	MOSTITLE	DISTANCE
1	11B	Infantryman	2.28685
2	16S	MANPADS Crewman	1.99142
3	19K	M1 ABRAMS Armor Crewman	2.98719
4	88M	Motor Transport Operator	3.23357

CLSNUM=2

OBS	MOS	MOSTITLE	DISTANCE
5	55B	Ammunition Specialist	3.10573
6	63B	Light Wheel Vehicle Mechanic	2.15722
7	67N	Utility Helicopter Repairer	2.24662

CLSNUM=3

OBS	MOS	MOSTITLE	DISTANCE
8	31C	Single Channel Radio Operator	1.83879
9	31D	Mobile Subscriber Equipment	1.83879

CLSNUM=4

OBS	MOS	MOSTITLE	DISTANCE
10	12B	Combat Engineer	2.26873
11	54B	Chemical Operations Specialist	2.26873

CLSNUM=5

OBS	MOS	MOSTITLE	DISTANCE
12	27E	TOW/DRAGON Repairer	2.48972
13	76Y	Unit Supply Specialist	2.48972

CLSNUM=6

OBS	MOS	MOSTITLE	DISTANCE
14	71L	Administrative Specialist	0

CLSNUM=7

OBS	MOS	MOSTITLE	DISTANCE
15	96B	Intelligence Analyst	0

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
THIRTEEN CLUSTER SOLUTION

---------------------------------- CLSNUM=8 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
16	29E	Radio Repairer	0

---------------------------------- CLSNUM=9 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
17	13B	Cannon Crewman	0

---------------------------------- CLSNUM=10 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
18	51B	Carpentry/Masonry Specialist	0

---------------------------------- CLSNUM=11 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
19	94B	Food Service Specialist	0

---------------------------------- CLSNUM=12 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
20	95B	Military Police	0

---------------------------------- CLSNUM=13 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
21	91A	Medical Specialist	0

Table 2.
WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
THIRTEEN CLUSTER SOLUTION
CLUSTER MEAN DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=1

N	Mean	Std Dev	Minimum	Maximum
4	2.6247570	0.5823268	1.9914173	3.2335681

CLSNUM=2

N	Mean	Std Dev	Minimum	Maximum
3	2.5031917	0.5237257	2.1572239	3.1057316

CLSNUM=3

N	Mean	Std Dev	Minimum	Maximum
2	1.8387945	0	1.8387945	1.8387945

CLSNUM=4

N	Mean	Std Dev	Minimum	Maximum
2	2.2687343	0	2.2687343	2.2687343

CLSNUM=5

N	Mean	Std Dev	Minimum	Maximum
2	2.4897170	0	2.4897170	2.4897170

CLSNUM=6

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
THIRTEEN CLUSTER SOLUTION
CLUSTER MEAN DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=7

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=8

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=9

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=10

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=11

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=12

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
THIRTEEN CLUSTER SOLUTION
CLUSTER MEAN DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

------------------------------- CLSNUM=13 -------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3

WARD HIERARCHICAL CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
THIRTEEN CLUSTER SOLUTION
CLUSTER MEAN PRINCIPAL COMPONENT SCORES

OBS	CLSNUM	FREQ	LGEDIST	PRIN1	PRIN2	PRIN3	PRIN4
1	1	4	3.23357	0.94178	-0.41443	-0.54779	-0.20067
2	2	3	3.10573	-0.42651	-0.33301	-0.14893	0.80895
3	3	2	1.83879	-0.03446	-0.54932	1.83055	-0.53762
4	4	2	2.26873	0.93563	-0.44990	-0.27634	-0.49252
5	5	2	2.48972	-0.75605	-0.13351	0.27045	0.26938
6	6	1	0.00000	-2.09002	0.85199	-0.87992	-0.97709
7	7	1	0.00000	-0.47021	1.16373	0.50459	-2.99176
8	8	1	0.00000	-1.02942	-0.45036	1.25971	0.52866
9	9	1	0.00000	1.39505	-0.53810	-0.53739	-0.03464
10	10	1	0.00000	-0.28777	-0.63798	-0.71623	1.30144
11	11	1	0.00000	-1.11918	-0.00989	-0.99476	0.50595
12	12	1	0.00000	1.19009	3.53967	1.00244	1.70296
13	13	1	0.00000	-0.36638	1.00311	-0.64980	-0.13818

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
1	0.52459	0.40278	-0.78229	-0.75384	-0.03248	-0.21080
2	0.18022	0.23969	0.36581	-0.31503	0.90909	0.04676
3	-0.35689	-1.42252	-0.35292	-0.34876	-0.72000	0.40141
4	-1.17451	-0.25019	0.25518	0.24890	1.03419	-0.75503
5	0.81277	0.07552	1.56654	0.66658	0.83485	-0.04589
6	0.29388	0.41377	0.26363	-0.25843	-1.74513	-1.02408
7	-0.77352	1.59005	0.30626	-0.20819	0.46813	0.96027
8	0.50260	0.26242	-0.29731	-0.16290	-0.82782	-2.27228
9	1.47176	-0.51123	1.27685	1.05951	-1.94432	1.34024
10	-3.02318	0.95869	0.34065	0.57528	-1.24062	1.11705
11	0.13162	-1.29561	-0.59487	-0.72441	0.09462	1.31918
12	-0.02328	0.32821	0.06542	-0.69082	-0.22590	0.21575
13	0.21834	-0.88210	-2.26651	3.23695	0.52561	-0.15420

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15
1	-0.31811	-0.03758	-0.84938	-0.04048	0.03439
2	1.06671	-0.70488	0.34592	-0.51229	1.12200
3	0.59410	0.86897	-0.26772	-0.39212	0.46121
4	0.00267	0.63155	1.51155	-0.27803	-0.94591
5	-0.65630	0.63200	-1.27951	0.16917	-0.89516
6	-0.13779	0.57816	0.41227	-2.06271	-0.27828
7	0.04356	-1.38691	0.12200	0.83647	0.16399
8	-0.02027	-0.65697	0.71673	2.33296	-0.11836
9	0.53475	-1.08848	0.98724	-0.18954	-1.13085
10	0.02439	0.32011	-1.01980	0.71867	-0.05276
11	-2.79041	0.07302	1.46694	0.90678	0.73551
12	0.02430	0.42866	0.28414	0.04179	-0.35312
13	0.51287	-0.26770	-0.53841	0.11638	0.29005

Appendix K

Average Linkage Hierarchical Cluster Analysis of Synthetic Validity CTI Data

- Plot of Cubic Clustering Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS

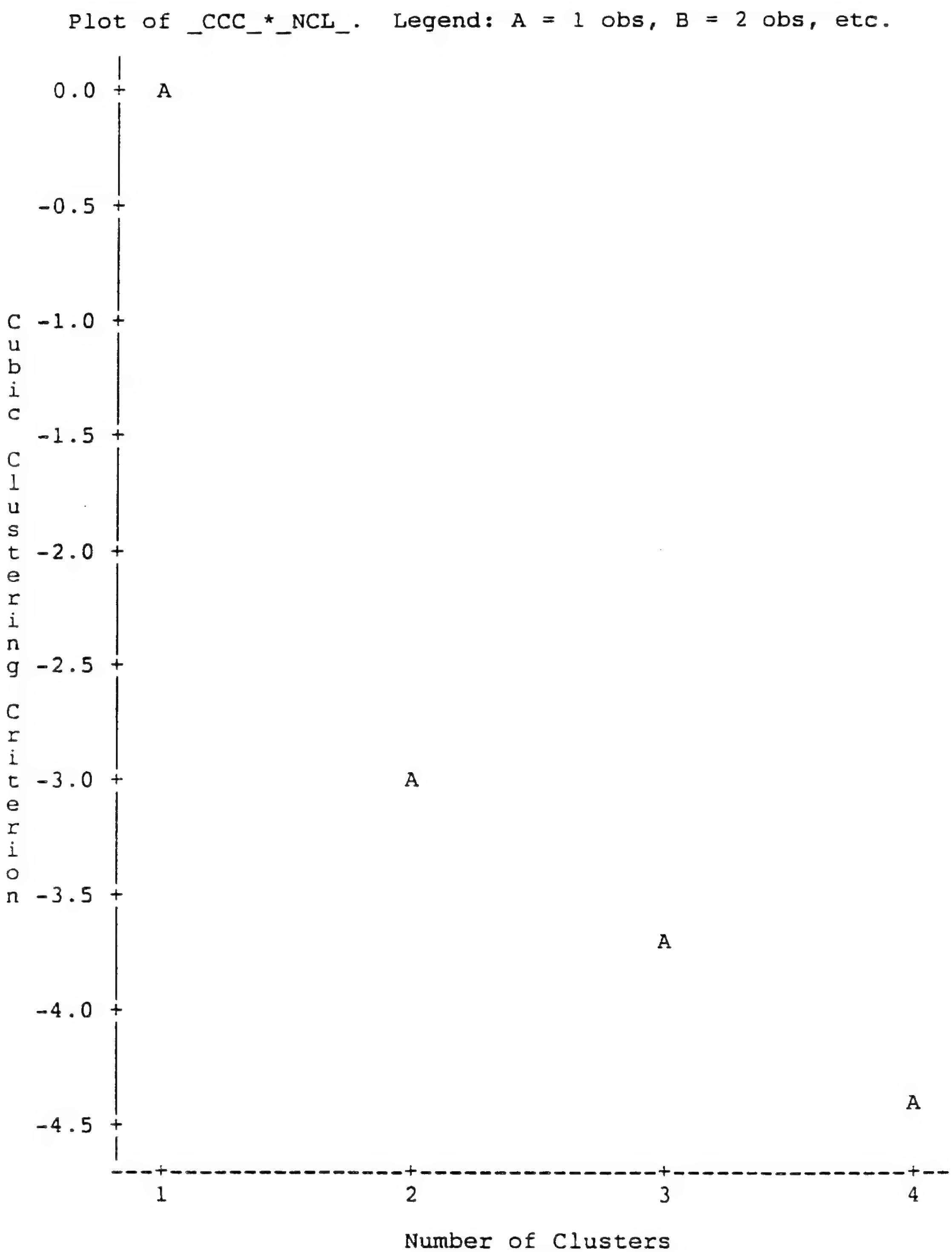


Figure 1. CCC by Number of Clusters

AVERAGE LINKAGE HIERARCHICAL CLUSTER ANALYSIS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

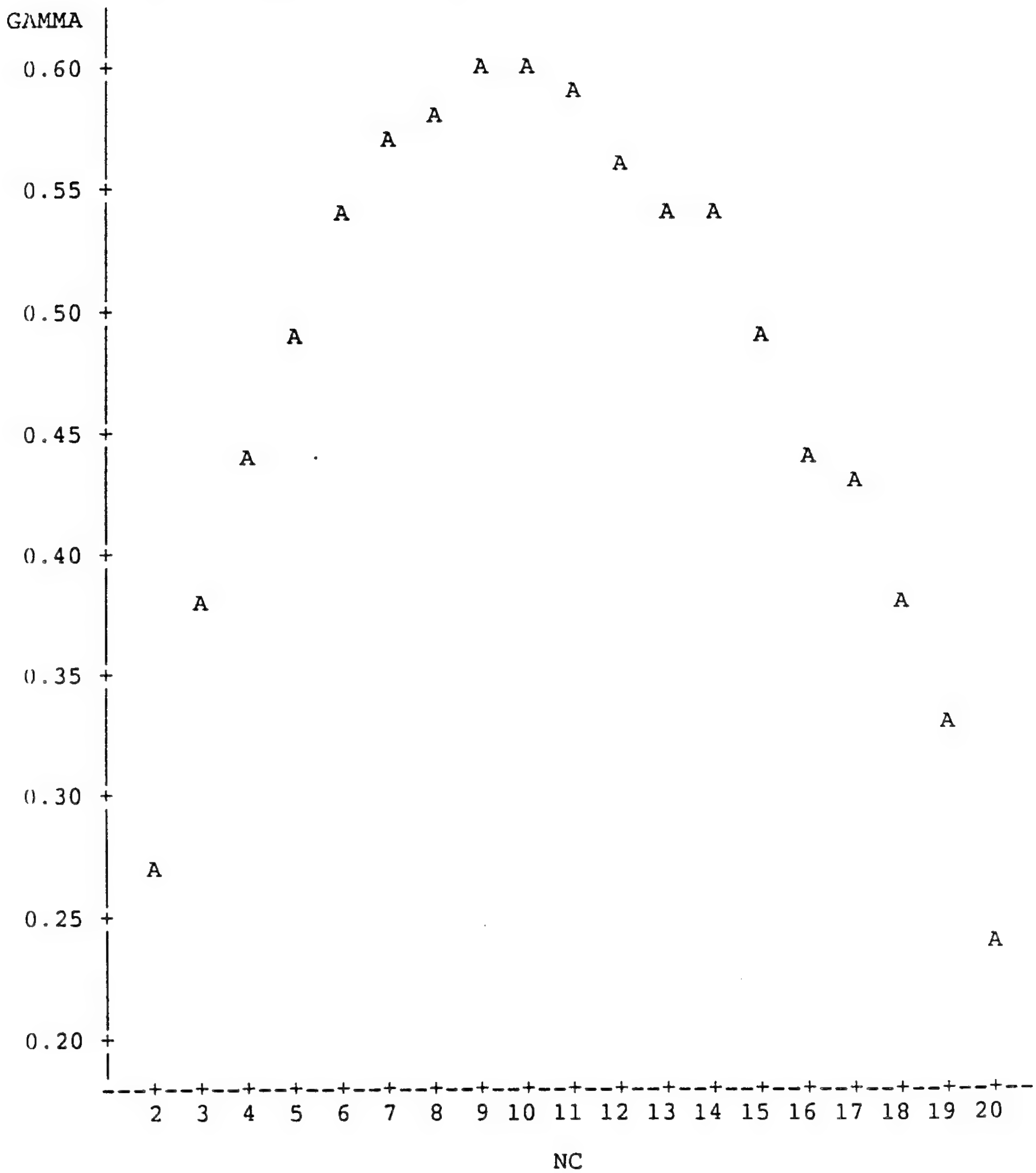


Figure 2. Gamma by Number of Clusters

Table 1.
AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
NINE CLUSTER SOLUTION

---------------------------------- CLSNUM=1 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
1	11B	Infantryman	2.76476
2	12B	Combat Engineer	3.66938
3	16S	MANPADS Crewman	2.80904
4	19K	M1 ABRAMS Armor Crewman	3.46615
5	27E	TOW/DRAGON Repairer	3.71489
6	31C	Single Channel Radio Operator	3.17915
7	31D	Mobile Subscriber Equipment	3.55096
8	54B	Chemical Operations Specialist	3.47017
9	63B	Light Wheel Vehicle Mechanic	2.67248
10	67N	Utility Helicopter Repairer	3.34570
11	76Y	Unit Supply Specialist	3.77987
12	88M	Motor Transport Operator	3.66332

---------------------------------- CLSNUM=2 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
13	29E	Radio Repairer	2.71079
14	55B	Ammunition Specialist	2.71079

---------------------------------- CLSNUM=3 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
15	71L	Administrative Specialist	0

---------------------------------- CLSNUM=4 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
16	13B	Cannon Crewman	0

---------------------------------- CLSNUM=5 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
17	51B	Carpentry/Masonry Specialist	0

---------------------------------- CLSNUM=6 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
18	96B	Intelligence Analyst	0

AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
NINE CLUSTER SOLUTION

---------------------------------- CLSNUM=7 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
19	94B	Food Service Specialist	0

---------------------------------- CLSNUM=8 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
20	95B	Military Police	0

---------------------------------- CLSNUM=9 ----------------------------------

OBS	MOS	MOSTITLE	DISTANCE
21	91A	Medical Specialist	0

Table 2

AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
NINE CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

CLSNUM=1

N	Mean	Std Dev	Minimum	Maximum
12	3.3404878	0.3941285	2.6724787	3.7798698

CLSNUM=2

N	Mean	Std Dev	Minimum	Maximum
2	2.7107930	0	2.7107930	2.7107930

CLSNUM=3

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=4

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=5

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

CLSNUM=6

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
NINE CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

---------------------------------- CLSNUM=7 ----------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

---------------------------------- CLSNUM=8 ----------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

---------------------------------- CLSNUM=9 ----------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3.

AVERAGE LINKAGE CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
NINE CLUSTER SOLUTION
MEAN CLUSTER PRINCIPAL COMPONENT SCORES

OBS	CLSNUM	FREQ	LGEDIST	PRIN1	PRIN2	PRIN3	PRIN4
1	1	12	3.77987	0.24502	-0.42860	0.16796	-0.02369
2	2	2	2.71079	-0.59590	-0.11468	0.12777	0.45779
3	3	1	0.00000	-2.09002	0.85199	-0.87992	-0.97709
4	4	1	0.00000	1.39505	-0.53810	-0.53739	-0.03464
5	5	1	0.00000	-0.28777	-0.63798	-0.71623	1.30144
6	6	1	0.00000	-0.47021	1.16373	0.50459	-2.99176
7	7	1	0.00000	-1.11918	-0.00989	-0.99476	0.50595
8	8	1	0.00000	1.19009	3.53967	1.00244	1.70296
9	9	1	0.00000	-0.36638	1.00311	-0.64980	-0.13818

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
1	0.10911	0.06021	-0.04139	-0.17559	0.35118	-0.09609
2	0.19751	-0.66214	0.55264	-0.44142	-0.07329	-1.31061
3	0.29388	0.41377	0.26363	-0.25843	-1.74513	-1.02408
4	1.47176	-0.51123	1.27685	1.05951	-1.94432	1.34024
5	-3.02318	0.95869	0.34065	0.57528	-1.24062	1.11705
6	-0.77352	1.59005	0.30626	-0.20819	0.46813	0.96027
7	0.13162	-1.29561	-0.59487	-0.72441	0.09462	1.31918
8	-0.02328	0.32821	0.06542	-0.69082	-0.22590	0.21575
9	0.21834	-0.88210	-2.26651	3.23695	0.52561	-0.15420

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15
1	0.01043	0.32463	-0.19307	-0.28174	-0.05228
2	0.83162	-1.27619	0.30121	1.50653	0.62639
3	-0.13779	0.57816	0.41227	-2.06271	-0.27828
4	0.53475	-1.08848	0.98724	-0.18954	-1.13085
5	0.02439	0.32011	-1.01980	0.71867	-0.05276
6	0.04356	-1.38691	0.12200	0.83647	0.16399
7	-2.79041	0.07302	1.46694	0.90678	0.73551
8	0.02430	0.42866	0.28414	0.04179	-0.35312
9	0.51287	-0.26770	-0.53841	0.11638	0.29005

Appendix L

K-Means Cluster Analysis of Synthetic Validity CTI Data

- Plot of Cubic Clustering Criterion by Number of Clusters
- Plot of Gamma Values by Number of Clusters
- Composition of Clusters
- Mean Cluster Distances
- Mean Cluster Factor Scores

SAS MODIFIED K-MEANS ITERATIVE CLUSTER ANALYSIS
CCC DIFFERENCES

Plot of _CCC_*_NCL_. Legend: A = 1 obs, B = 2 obs, etc.

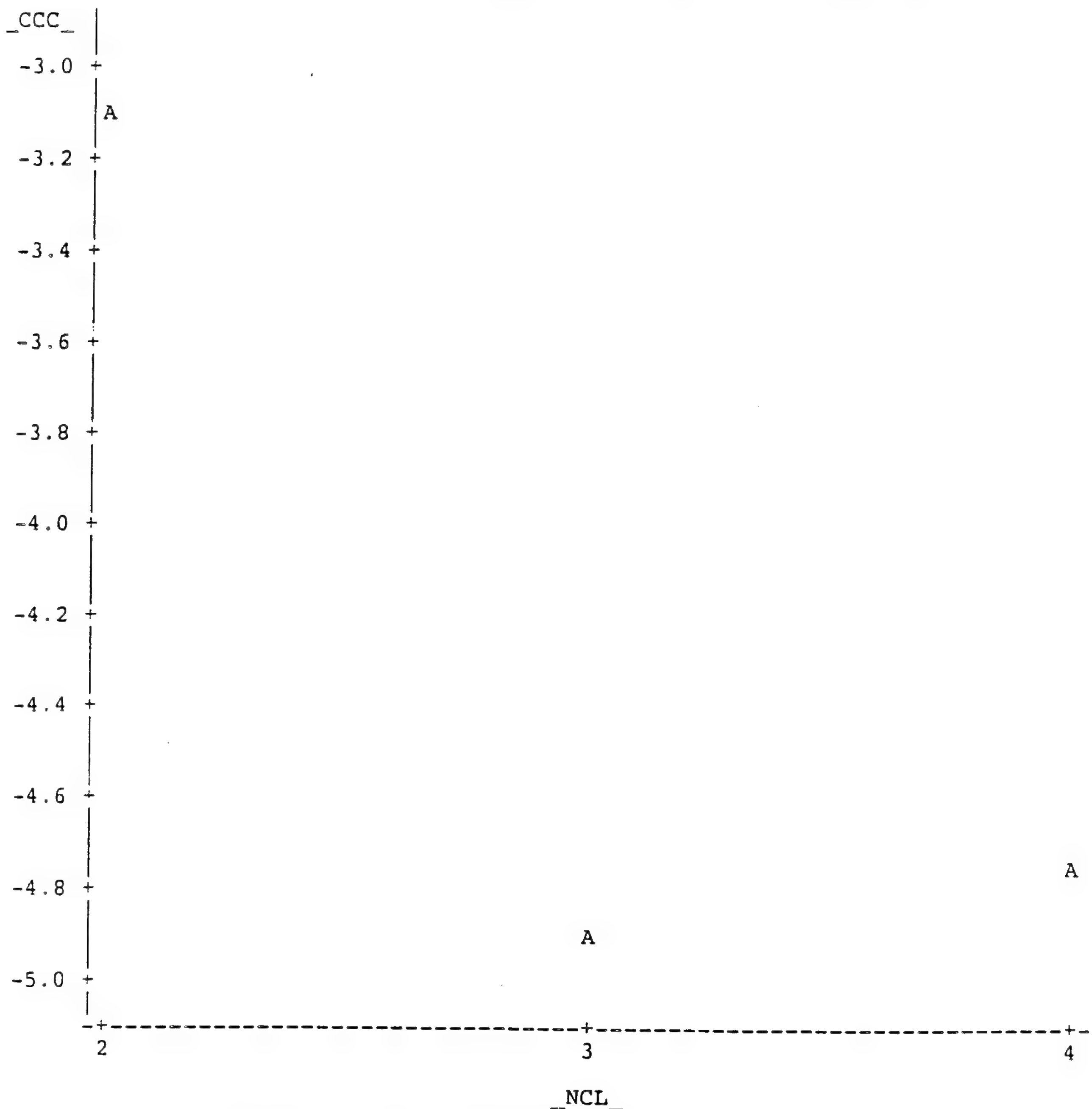


Figure 1. CCC by Number of Clusters

SAS MODIFIED K-MEANS ITERATIVE CLUSTER ANALYSIS

Plot of GAMMA*NC. Legend: A = 1 obs, B = 2 obs, etc.

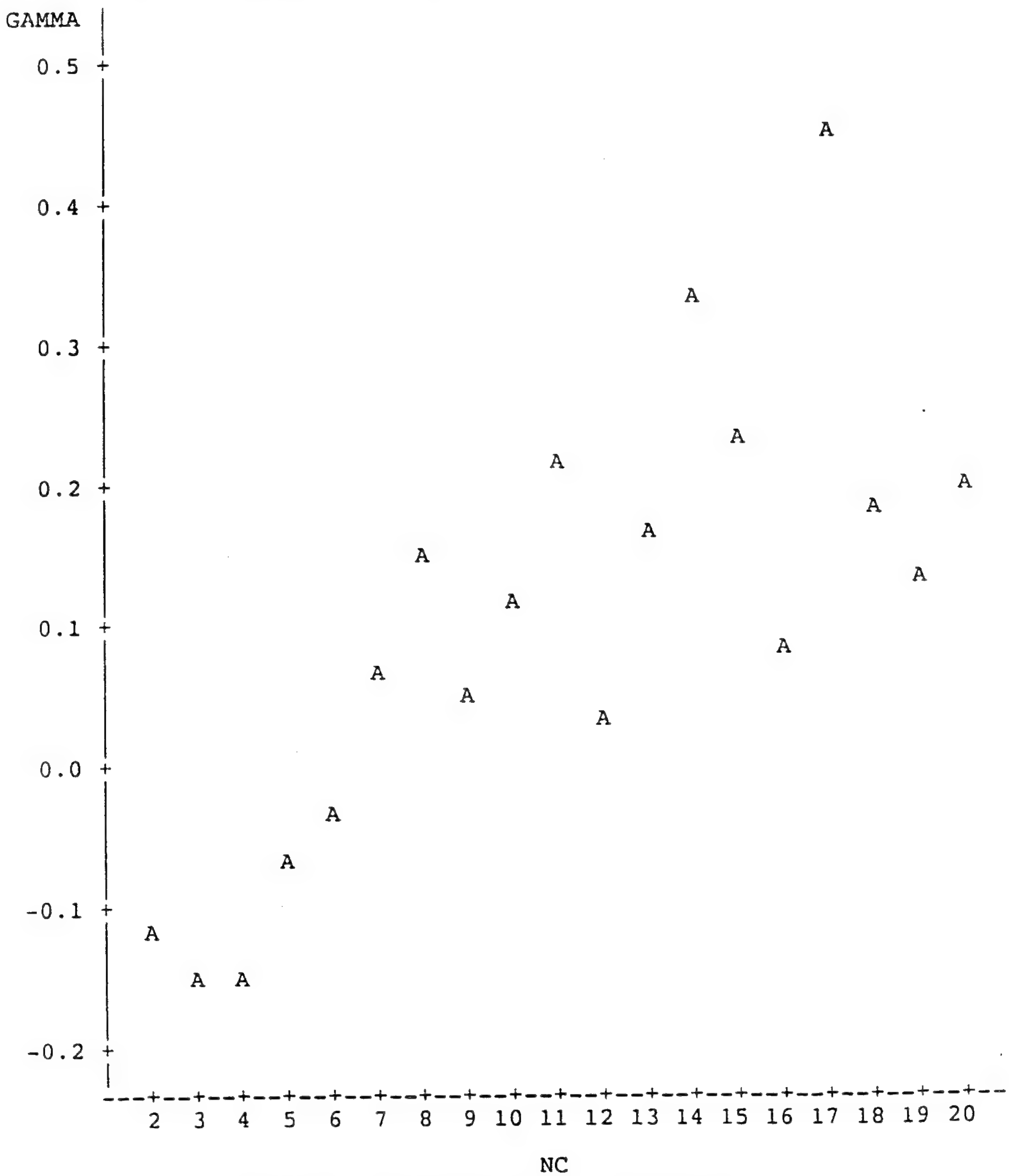


Figure 2. Gamma by Number of Clusters

Table 1.
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION

Cluster=1

OBS	MOS	MOSTITLE	DISTANCE
1	11B	Infantryman	1.50805
2	16S	MANPADS Crewman	1.50805

Cluster=2

OBS	MOS	MOSTITLE	DISTANCE
3	12B	Combat Engineer	0

Cluster=3

OBS	MOS	MOSTITLE	DISTANCE
4	13B	Cannon Crewman	0

Cluster=4

OBS	MOS	MOSTITLE	DISTANCE
5	94B	Food Service Specialist	0

Cluster=5

OBS	MOS	MOSTITLE	DISTANCE
6	19K	M1 ABRAMS Armor Crewman	0

Cluster=6

OBS	MOS	MOSTITLE	DISTANCE
7	27E	TOW/DRAGON Repairer	0

Cluster=7

OBS	MOS	MOSTITLE	DISTANCE
8	29E	Radio Repairer	0

Cluster=8

OBS	MOS	MOSTITLE	DISTANCE
9	95B	Military Police	0

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION

------------------------------ Cluster=9 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
10	31D	Mobile Subscriber Equipment	0

------------------------------ Cluster=10 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
11	51B	Carpentry/Masonry Specialist	0

------------------------------ Cluster=11 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
12	31C	Single Channel Radio Operator	1.64476
13	54B	Chemical Operations Specialist	1.64476

------------------------------ Cluster=12 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
14	55B	Ammunition Specialist	0

------------------------------ Cluster=13 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
15	88M	Motor Transport Operator	2.39230
16	63B	Light Wheel Vehicle Mechanic	2.50947
17	96B	Intelligence Analyst	3.47328

------------------------------ Cluster=14 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
18	67N	Utility Helicopter Repairer	0

------------------------------ Cluster=15 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
19	71L	Administrative Specialist	0

------------------------------ Cluster=16 ------------------------------

OBS	MOS	MOSTITLE	DISTANCE
20	76Y	Unit Supply Specialist	0

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION

------------------------------- Cluster=17 ---------------------------------

OBS	MOS	MOSTITLE	DISTANCE
21	91A	Medical Specialist	0

Table 2.
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

Cluster=1

N	Mean	Std Dev	Minimum	Maximum
2	1.5080474	0	1.5080474	1.5080474

Cluster=2

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=3

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=4

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=5

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=6

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

Cluster=7

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=8

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=9

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=10

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Cluster=11

N	Mean	Std Dev	Minimum	Maximum
2	1.6447628	0	1.6447628	1.6447628

Cluster=12

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION
MEAN CLUSTER DISTANCES

Analysis Variable : DISTANCE Distance to Cluster Seed

------------------------------ Cluster=13 ------------------------------

N	Mean	Std Dev	Minimum	Maximum
3	2.7916829	0.5931787	2.3922987	3.4732776

------------------------------ Cluster=14 ------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

------------------------------ Cluster=15 ------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

------------------------------ Cluster=16 ------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

------------------------------ Cluster=17 ------------------------------

N	Mean	Std Dev	Minimum	Maximum
1	0	.	0	0

Table 3.
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION
MEAN CLUSTER PRINCIPAL COMPONENT SCORES

OBS	FREQ	LGEDIST	NEAREST	PRIN1	PRIN2	PRIN3	PRIN4
1	2	1.50805	13	1.06559	-0.02192	-0.45427	-0.57613
2	1	0.00000	1	1.76459	-0.08171	-0.31079	-0.68709
3	1	0.00000	13	1.35409	-0.24669	-0.70428	-0.33774
4	1	0.00000	13	-1.45349	-0.30776	-0.82867	0.13695
5	1	0.00000	1	0.87978	-0.46008	-1.09979	-0.14361
6	1	0.00000	14	0.06958	-0.68118	1.10102	1.14129
7	1	0.00000	13	-0.40553	-0.51579	1.06233	0.91681
8	1	0.00000	13	0.29742	4.01766	0.31370	1.35647
9	1	0.00000	14	-0.31905	-0.61730	2.07215	0.80407
10	1	0.00000	13	0.48131	-0.40400	-0.87607	1.15929
11	2	1.64476	13	0.65857	-0.49530	0.93073	-0.29245
12	1	0.00000	13	-0.67466	0.03274	-0.79051	-0.22871
13	3	3.47328	14	-0.36026	-0.05751	-0.13307	-0.36114
14	1	0.00000	13	-0.34172	-0.55708	0.30183	1.25862
15	1	0.00000	13	-2.02128	0.41597	-0.09448	-1.38610
16	1	0.00000	12	-1.72717	-0.09264	-0.56586	-0.66921

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
1	0.42507	-0.32733	-0.78898	-0.32759	-0.52455	-0.74904
2	-1.74661	0.83895	0.74106	-0.27315	-1.22547	-1.21985
3	1.51465	2.00153	0.19689	-0.22295	1.99702	0.55974
4	-0.34233	0.80503	-0.92621	0.16350	-0.38058	-0.05036
5	1.62599	-1.63734	-0.21287	-0.45958	1.24708	0.16719
6	0.84554	0.46725	2.02812	-0.38311	-0.40622	0.29944
7	0.49292	0.08575	0.42721	-0.37540	-0.76070	-1.68435
8	0.17154	0.03094	-0.30882	-0.57616	-0.10378	0.25632
9	-0.29532	0.34938	-2.25117	-0.05229	1.39465	-0.36091
10	-3.04245	-0.76095	0.62027	-0.24751	2.00416	0.28422
11	-0.11382	0.08625	-0.34777	0.41023	-0.78620	0.31054
12	-0.56347	1.97140	-0.29016	-0.52110	-0.89249	0.89679
13	0.01580	-1.25014	-0.25665	0.12486	-0.52617	0.79857
14	0.32534	-0.51443	0.96856	-0.04972	-0.35425	1.14074
15	-0.02133	-0.59641	0.58385	-0.94271	0.88692	-2.30650
16	0.24122	0.96394	0.74618	-0.63031	0.23631	1.00246

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15
1	0.21418	-1.47768	-0.45476	0.69564	-0.75838
2	-0.03760	0.58639	-1.33608	-0.21497	1.70620
3	-0.09582	0.43744	0.35718	-2.06825	0.02871
4	3.43055	-0.34337	0.57864	-0.67882	0.55796
5	0.00772	0.45125	0.53384	1.56368	0.89599
6	0.61493	-1.68962	-0.29945	0.18506	-1.45998
7	-0.87886	-0.80742	1.89526	-0.67812	1.77516
8	0.22770	0.25194	0.23404	0.17302	0.07488
9	-0.33737	0.27224	-2.04015	0.06457	0.29450
10	0.00911	-0.50982	0.95462	0.24195	-0.52744
11	0.44787	1.77061	1.14309	0.68382	-0.94665
12	-2.04973	-0.12379	0.41826	0.26454	-1.00536
13	-0.52884	-0.08015	0.03722	-1.15145	0.39294
14	0.22637	0.59605	-1.77801	-0.27759	-0.09764
15	-0.34338	0.88374	-0.23389	-0.52184	-1.62093
16	-0.17610	0.14298	-0.47824	2.27991	1.57933

Table 3.
SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION
MEAN CLUSTER PRINCIPAL COMPONENT SCORES

OBS	FREQ	LGEDIST	NEAREST	PRIN1	PRIN2	PRIN3	PRIN4
1	2	1.50805	13	1.06559	-0.02192	-0.45427	-0.57613
2	1	0.00000	1	1.76459	-0.08171	-0.31079	-0.68709
3	1	0.00000	13	1.35409	-0.24669	-0.70428	-0.33774
4	1	0.00000	13	-1.45349	-0.30776	-0.82867	0.13695
5	1	0.00000	1	0.87978	-0.46008	-1.09979	-0.14361
6	1	0.00000	14	0.06958	-0.68118	1.10102	1.14129
7	1	0.00000	13	-0.40553	-0.51579	1.06233	0.91681
8	1	0.00000	13	0.29742	4.01766	0.31370	1.35647
9	1	0.00000	14	-0.31905	-0.61730	2.07215	0.80407
10	1	0.00000	13	0.48131	-0.40400	-0.87607	1.15929
11	2	1.64476	13	0.65857	-0.49530	0.93073	-0.29245
12	1	0.00000	13	-0.67466	0.03274	-0.79051	-0.22871
13	3	3.47328	14	-0.36026	-0.05751	-0.13307	-0.36114
14	1	0.00000	13	-0.34172	-0.55708	0.30183	1.25862
15	1	0.00000	13	-2.02128	0.41597	-0.09448	-1.38610
16	1	0.00000	12	-1.72717	-0.09264	-0.56586	-0.66921

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
1	0.42507	-0.32733	-0.78898	-0.32759	-0.52455	-0.74904
2	-1.74661	0.83895	0.74106	-0.27315	-1.22547	-1.21985
3	1.51465	2.00153	0.19689	-0.22295	1.99702	0.55974
4	-0.34233	0.80503	-0.92621	0.16350	-0.38058	-0.05036
5	1.62599	-1.63734	-0.21287	-0.45958	1.24708	0.16719
6	0.84554	0.46725	2.02812	-0.38311	-0.40622	0.29944
7	0.49292	0.08575	0.42721	-0.37540	-0.76070	-1.68435
8	0.17154	0.03094	-0.30882	-0.57616	-0.10378	0.25632
9	-0.29532	0.34938	-2.25117	-0.05229	1.39465	-0.36091
10	-3.04245	-0.76095	0.62027	-0.24751	2.00416	0.28422
11	-0.11382	0.08625	-0.34777	0.41023	-0.78620	0.31054
12	-0.56347	1.97140	-0.29016	-0.52110	-0.89249	0.89679
13	0.01580	-1.25014	-0.25665	0.12486	-0.52617	0.79857
14	0.32534	-0.51443	0.96856	-0.04972	-0.35425	1.14074
15	-0.02133	-0.59641	0.58385	-0.94271	0.88692	-2.30650
16	0.24122	0.96394	0.74618	-0.63031	0.23631	1.00246

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15
1	0.21418	-1.47768	-0.45476	0.69564	-0.75838
2	-0.03760	0.58639	-1.33608	-0.21497	1.70620
3	-0.09582	0.43744	0.35718	-2.06825	0.02871
4	3.43055	-0.34337	0.57864	-0.67882	0.55796
5	0.00772	0.45125	0.53384	1.56368	0.89599
6	0.61493	-1.68962	-0.29945	0.18506	-1.45998
7	-0.87886	-0.80742	1.89526	-0.67812	1.77516
8	0.22770	0.25194	0.23404	0.17302	0.07488
9	-0.33737	0.27224	-2.04015	0.06457	0.29450
10	0.00911	-0.50982	0.95462	0.24195	-0.52744
11	0.44787	1.77061	1.14309	0.68382	-0.94665
12	-2.04973	-0.12379	0.41826	0.26454	-1.00536
13	-0.52884	-0.08015	0.03722	-1.15145	0.39294
14	0.22637	0.59605	-1.77801	-0.27759	-0.09764
15	-0.34338	0.88374	-0.23389	-0.52184	-1.62093
16	-0.17610	0.14298	-0.47824	2.27991	1.57933

SAS K-MEANS CLUSTER ANALYSIS OF SYNTHETIC VALIDITY CTI DATA
SEVENTEEN CLUSTER SOLUTION
MEAN CLUSTER PRINCIPAL COMPONENT SCORES

OBS	FREQ	LGEDIST	NEAREST	PRIN1	PRIN2	PRIN3	PRIN4
17	1	0.00000	13	-0.27141	0.70482	-0.13430	-0.50050

OBS	PRIN5	PRIN6	PRIN7	PRIN8	PRIN9	PRIN10
17	0.12441	0.22752	0.72054	4.03065	0.55734	-0.50364

OBS	PRIN11	PRIN12	PRIN13	PRIN14	PRIN15
17	-0.33510	-0.49346	-0.29432	0.36233	0.02985